

# SEARCH REQUEST FORM

Requestor's Name: \_\_\_\_\_ Serial Number: \_\_\_\_\_

Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

## STAFF USE ONLY

Date completed: 02-10-03  
Searcher: Beverly C 4999  
Terminal time: 25  
Elapsed time: \_\_\_\_\_  
CPU time: \_\_\_\_\_  
Total time: 25  
Number of Searches: \_\_\_\_\_  
Number of Databases: 1

### Search Site

\_\_\_\_\_ STIC  
\_\_\_\_\_ CM-1  
\_\_\_\_\_ Pre-S

### Type of Search

\_\_\_\_\_ N.A. Sequence  
\_\_\_\_\_ A.A. Sequence  
\_\_\_\_\_ Structure  
\_\_\_\_\_ Bibliographic

### Vendors

\_\_\_\_\_ IG Suite  
\_\_\_\_\_ STN  
\_\_\_\_\_ Dialog  
\_\_\_\_\_ APS  
\_\_\_\_\_ Geninfo  
\_\_\_\_\_ SDC  
\_\_\_\_\_ DARC/Questel  
\_\_\_\_\_ Other: CGN



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2003, 08:35:55 ; Search time 2704 Seconds

(without alignments)  
9288.358 Million cell updates/sec

Title: US-10-021-811-35

Sequence: 1 ggcagagctctatcacacac.....aaaaaaaaaaaaaaaaa 863

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_hcg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_hcg\_hum: \*  
31: em\_hcg\_inv: \*  
32: em\_hcg\_other: \*  
33: em\_hcg\_mus: \*  
34: em\_hcg\_pln: \*  
35: em\_hcg\_rnd: \*  
36: em\_hcg\_mam: \*  
37: em\_hcg\_vrt: \*  
38: em\_sy: \*  
39: em\_hcg\_hum: \*  
40: em\_hcg\_mus: \*  
41: em\_hcg\_other: \*

score greater than or equal to the score of the result being printed,  
and its derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	4.9	52514	2 AC116551	AC116551 Dictyoste
2	41	4.8	1455	9 BC009883	BC009883 Homo sapi
3	41	4.8	2037	6 AX317981	AX317981 Sequence
4	41	4.8	2097	3 AB079606	AB079606 Dictyoste
5	41	4.8	2307	6 AB1160	AB1160 Sequence 5
6	41	4.8	4329	6 AY1162	AY1162 Sequence 7
7	40	4.6	1856	3 AY051514	AY051514 Drosophi
8	40	4.6	2472	9 BC025672	BC025672 Homo sapi
9	40	4.6	10825	6 AX344799	AX344799 Sequence
10	39	4.5	200	11 G38015	G38015 C1M31 plasm
11	39	4.5	236	11 G37836	G37836 wh_hspi pla
12	39	4.5	415	8 AB072484	AB072484 Lycopersi
13	39	4.5	691	3 AY070797	AY070797 Drosophi
14	39	4.5	988	10 BC028785	BC028785 Mus muscu
15	39	4.5	1121	9 AF236637	AF236637 Homo sapi
16	39	4.5	1373	8 AY124468	AY124468 Hevea bra
17	39	4.5	1470	9 BC012575	BC012575 Homo sapi
18	39	4.5	2547	3 PFGRPXA	PFGRPXA P. italicu
19	39	4.5	2554	3 AF106064	AF106064 Plasmodi
20	39	4.5	3001	6 AX344925	AX344925 Sequence
21	39	4.5	5273	6 AX345777	AX345777 Sequence
22	39	4.5	5313	6 AX345665	AX345665 Sequence
23	39	4.5	6120	6 AX323825	AX323825 Sequence
24	39	4.5	6120	6 AX344894	AX344894 Sequence
25	39	4.5	6476	6 HSM803020	AL171373 Homo sapi
26	39	4.5	6956	6 AX348657	AX348657 Sequence
27	39	4.5	7046	6 AX344880	AX344880 Sequence
28	39	4.5	7890	6 AX251883	AX347018 Sequence
29	39	4.5	7890	6 AX251883	AX251883 Sequence
30	39	4.5	8246	6 AX345105	AX348672 Sequence
31	39	4.5	8246	6 AX345107	AX345105 Sequence
32	39	4.5	8372	6 AX251107	AX251107 Sequence
33	39	4.5	9963	6 AX345596	AX345596 Sequence
34	39	4.5	11178	6 AX344168	AX251754 Sequence
35	39	4.5	11178	6 AX344168	AX344168 Sequence
36	39	4.5	11178	6 AX348939	AX348939 Sequence
37	39	4.5	13427	6 AX346828	AX346828 Sequence
38	39	4.5	14923	3 AE001394	AE001394 Plasmodi
39	39	4.5	14955	3 AF000580	AF000580 Dictyoste
40	39	4.5	14955	3 AF000580	AF000580 Dictyoste
41	39	4.5	15500	6 AX251262	AX251262 Sequence
42	39	4.5	17234	6 AX458502	AX458502 Sequence
43	39	4.5	17389	6 AX346316	AX346316 Sequence
44	39	4.5	35119	2 AC116958	AC116958 Dictyoste
45	39	4.5	80472	8 F24J2	AF262039 Arabidops
46	39	4.5	96136	2 AC116548	AC116548 Dictyoste
47	39	4.5	97683	2 AC116548	AC116548 Dictyoste
48	39	4.5	98734	2 PFMALIP2	AL031745 Plasmodi
49	39	4.5	101567	2 AL845510	AL845510 Dario ter
50	39	4.5	113880	3 PFMAL3P4	AL008970 Plasmodi
51	39	4.5	116696	3 PFMAL3P3	Z98547 Plasmodi
52	39	4.5	137526	2 AC125902	AC125902 Rattus no
53	39	4.5	205429	2 AC005506	AC005506 Plasmodi
54	39	4.5	205429	2 AC005506	AC005506 Plasmodi
55	39	4.5	216100	2 AC129180	AC129180 Mus muscu
56	39	4.5	232519	2 AC095709	AC095709 Rattus no
57	39	4.5	234112	3 PFMAL4P2	AL035475 Plasmodi
58	39	4.5	349980	6 AX344566	AX344566 Sequence
59	39	4.5	349980	6 AX344567	AX344567 Sequence
60	38	4.4	501	8 AF048745	AF048745 Phytolacc
61	38	4.4	609	3 TRBSLEXA	X00632 T. cruzi sma
62	38	4.4	1588	3 AY069348	AY069348 Drosophi
63	38	4.4	2120	9 BC008496	BC008496 Homo sapi
64	38	4.4	2120	9 BC008496	BC008496 Homo sapi
65	38	4.4	2682	9 BC014993	BC014993 Homo sapi

66	38	4.4	4197	3	AF163835	AF163835 Dictyoste	C 139	36	4.2	314	11	G293110	G293110 human SRS E
67	38	4.4	4446	3	DDU72236	DDU72236 Dictyosteli	C 140	36	4.2	332	6	AX187103	AX187103 Sequence
68	38	4.4	5447	6	AX251515	AX251515 Sequence	C 141	36	4.2	414	6	AF15979	AF15979 Butrus ma
69	38	4.4	6386	9	HSMB803518	HSMB803518 Homo sapi	C 142	36	4.2	500	11	G33488	G33488 human SRS S
70	38	4.4	7133	12	CUPYBULDBX	CUPYBULDBX Homo sapi	C 143	36	4.2	513	11	G29359	G29359 human SRS S
71	38	4.4	7805	8	YSCDYBULDBX	YSCDYBULDBX Homo sapi	C 144	36	4.2	664	9	HSAB27335	HSAB27335 Homo sapi
72	38	4.4	9770	6	AX344934	AX344934 Sequence	C 145	36	4.2	778	9	AK026857	AK026857 Homo sapi
73	38	4.4	17480	3	AC114258	AC114258 Dictyoste	C 146	36	4.2	838	9	AY061350	AY061350 Drosophila
74	38	4.4	19798	6	AX164588	AX164588 Sequence	C 147	36	4.2	840	8	CNS0180K	CNS0180K Homo sapi
75	38	4.4	19912	6	AX164586	AX164586 Sequence	C 148	36	4.2	840	8	CNS0180K	CNS0180K Homo sapi
76	38	4.4	20160	6	AX164590	AX164590 Sequence	C 149	36	4.2	963	3	AY118696	AY118696 Drosophila
77	38	4.4	20217	6	AX164594	AX164594 Sequence	C 150	36	4.2	1007	6	AR205208	AR205208 Sequence
78	38	4.4	20247	6	AX164596	AX164596 Sequence	C 151	36	4.2	1007	6	AR205209	AR205209 Sequence
79	38	4.4	20316	6	AX164592	AX164592 Sequence	C 152	36	4.2	1033	3	AY128476	AY128476 Drosophila
80	38	4.4	104992	2	AC005504	AC005504 Plasmodiu	C 153	36	4.2	1062	9	BC004155	BC004155 Homo sapi
81	38	4.4	141226	2	AC004710	AC004710 Plasmodiu	C 154	36	4.2	1072	10	BC008319	BC008319 Homo sapi
82	38	4.4	142160	2	AL353719	AL353719 Human DNA	C 155	36	4.2	1097	10	BC008162	BC008162 Mus muscu
83	38	4.4	169546	2	AC004157	AC004157 Plasmodiu	C 156	36	4.2	1127	9	HSMB801270	HSMB801270 Homo sapi
84	38	4.4	171187	2	AC116960	AC116960 Dictyoste	C 157	36	4.2	1127	9	BC031345	BC031345 Homo sapi
85	38	4.4	178273	2	AC005308	AC005308 Plasmodiu	C 158	36	4.2	1138	9	HSMB802391	HSMB802391 Homo sapi
86	38	4.4	196490	2	AC005507	AC005507 Plasmodiu	C 159	36	4.2	1153	9	AK027115	AK027115 Homo sapi
87	37	4.3	441	11	G73684	G73684 RZ976R etio	C 160	36	4.2	1176	9	IR0783647	IR0783647 Homo sapi
88	37	4.3	819	3	AF246695	AF246695 Bombyx mo	C 161	36	4.2	1193	9	BC004365	BC004365 Homo sapi
89	37	4.3	900	3	BC030229	BC030229 Homo sapi	C 162	36	4.2	1214	3	HSMB801955	HSMB801955 Homo sapi
90	37	4.3	1137	9	HSUB6818	HSUB6818 Human ubiq	C 163	36	4.2	1222	3	AY075185	AY075185 Homo sapi
91	37	4.3	1168	9	BC005085	BC005085 Homo sapi	C 164	36	4.2	1232	3	HSMB801798	HSMB801798 Homo sapi
92	37	4.3	1206	8	ATRNMX	ATRNMX Homo sapi	C 165	36	4.2	1268	8	AF48505	AF48505 Brassaica
93	37	4.3	1260	9	BC007235	BC007235 Homo sapi	C 166	36	4.2	1270	8	AF418694	AF418694 Drosophila
94	37	4.3	1261	3	DDIADCVAD01	DDIADCVAD01 Homo sapi	C 167	36	4.2	1279	6	AX086727	AX086727 Sequence
95	37	4.3	1492	8	HAU91341	HAU91341 Helianthus	C 168	36	4.2	1321	3	AY119217	AY119217 Sequence
96	37	4.3	1525	8	AF251264	AF251264 Trilicium	C 169	36	4.2	1384	9	AF193053	AF193053 Homo sapi
97	37	4.3	1700	8	AF321556	AF321556 Trilicium	C 170	36	4.2	1419	8	CCCHSMR	CCCHSMR Homo sapi
98	37	4.3	1718	9	AK025339	AK025339 Homo sapi	C 171	36	4.2	1428	3	DMRSSEMR	DMRSSEMR Homo sapi
99	37	4.3	1759	6	AX100593	AX100593 Sequence	C 172	36	4.2	1452	3	AF273828	AF273828 D. melanog
100	37	4.3	1856	10	AF394451	AF394451 Mus muscu	C 173	36	4.2	1454	10	BC028987	BC028987 Mus muscu
101	37	4.3	1856	9	HSUB6867	HSUB6867 Human ubiq	C 174	36	4.2	1501	3	BC015453	BC015453 Homo sapi
102	37	4.3	1867	10	BC005637	BC005637 Mus muscu	C 175	36	4.2	1542	3	AY070834	AY070834 Drosophila
103	37	4.3	1969	5	AY036624	AY036624 Dario rex	C 176	36	4.2	1551	10	MMU291757	MMU291757 Homo sapi
104	37	4.3	2008	3	AF145689	AF145689 Drosophila	C 177	36	4.2	1600	6	BC032456	BC032456 Homo sapi
105	37	4.3	2290	3	AY119138	AY119138 Drosophila	C 178	36	4.2	1614	9	BC022483	BC022483 Homo sapi
106	37	4.3	2441	3	DDU67940	DDU67940 Dictyosteli	C 179	36	4.2	1616	10	BC021155	BC021155 Mus muscu
107	37	4.3	2832	4	AF288406	AF288406 Homo sapi	C 180	36	4.2	1631	6	AX036060	AX036060 Sequence
108	37	4.3	3368	5	AF051786	AF051786 Xenopus 1	C 181	36	4.2	1654	3	AF503909	AF503909 Homo sapi
109	37	4.3	9750	12	AF025397	AF025397 Expressio	C 182	36	4.2	1736	3	AR119632	AR119632 Sequence
110	37	4.3	9760	12	EVU68815	EVU68815 Expression	C 183	36	4.2	1736	6	AR119632	AR119632 Sequence
111	37	4.3	9763	12	AF126280	AF126280 Expressio	C 184	36	4.2	1736	6	AR119632	AR119632 Sequence
112	37	4.3	10136	6	135495	135495 Sequence 2	C 185	36	4.2	1736	6	AR169113	AR169113 Sequence
113	37	4.3	10765	12	EVU67875	EVU67875 pBSP-1 year	C 186	36	4.2	1736	6	AR169113	AR169113 Sequence
114	37	4.3	12029	3	AE001381	AE001381 Plasmodiu	C 187	36	4.2	1865	10	BC016526	BC016526 Mus muscu
115	37	4.3	27785	2	AC116978	AC116978 Dictyoste	C 188	36	4.2	1940	10	BC026886	BC026886 Homo sapi
116	37	4.3	45296	2	AC116978	AC116978 Dictyoste	C 189	36	4.2	1958	3	AF242200	AF242200 Bombyx mo
117	37	4.3	56152	2	AC116963	AC116963 Dictyoste	C 190	36	4.2	1962	3	AY060994	AY060994 Drosophila
118	37	4.3	56152	6	AX251552	AX251552 Sequence	C 191	36	4.2	1973	9	AB063088	AB063088 Macaca fa
119	37	4.3	84759	2	AP000730	AP000730 Homo sapi	C 192	36	4.2	2096	9	BC022483	BC022483 Homo sapi
120	37	4.3	84805	2	AC116918	AC116918 Dictyoste	C 193	36	4.2	2130	9	BC025751	BC025751 Homo sapi
121	37	4.3	106144	2	AC116425	AC116425 Dictyoste	C 194	36	4.2	2248	9	AK025239	AK025239 Homo sapi
122	37	4.3	110000	2	AC0984053_1	Continuation (2 of	C 195	36	4.2	2253	3	BC032682	BC032682 Homo sapi
123	37	4.3	110000	2	AC098328_1	Continuation (2 of	C 196	36	4.2	2301	5	AF045607	AF045607 Xenopus 1
124	37	4.3	120240	2	AP000627	AP000627 Homo sapi	C 197	36	4.2	2318	3	DDU67940	DDU67940 Homo sapi
125	37	4.3	133501	2	AC116956	AC116956 Dictyoste	C 198	36	4.2	2349	3	DDU67940	DDU67940 Homo sapi
126	37	4.3	164599	2	AC068971	AC068971 Homo sapi	C 199	36	4.2	2394	9	HSMB805513	HSMB805513 Homo sapi
127	37	4.3	166125	2	AC019327	AC019327 Homo sapi	C 200	36	4.2	2427	9	AK000310	AK000310 Homo sapi
128	37	4.3	167357	2	AC095571	AC095571 Homo sapi	C 201	36	4.2	2430	9	AK074464	AK074464 Homo sapi
129	37	4.3	169547	2	AC095580	AC095580 Rattus no	C 202	36	4.2	2450	5	AF287007	AF287007 Dario rex
130	37	4.3	170001	2	AC116086	AC116086 Rattus no	C 203	36	4.2	2510	6	BC014039	BC014039 Homo sapi
131	37	4.3	179786	2	AP001015	AP001015 Homo sapi	C 204	36	4.2	2531	4	BC027040	BC027040 Homo sapi
132	37	4.3	196827	2	AC107825	AC107825 Mus muscu	C 205	36	4.2	2714	4	BTU75304	BTU75304 Bos taurus
133	37	4.3	202290	2	AL391335	AL391335 Homo sapi	C 206	36	4.2	2720	9	BC007575	BC007575 Homo sapi
134	37	4.3	217726	2	AC127573	AC127573 Mus muscu	C 207	36	4.2	2927	3	AY118573	AY118573 Homo sapi
135	37	4.3	253751	2	AC051621	AC051621 Mus muscu	C 208	36	4.2	3065	14	LRVBARCD	LRVBARCD Homo sapi
136	37	4.3	259854	2	AC098462	AC098462 Rattus no	C 209	36	4.2	3196	9	AF355402	AF355402 Homo sapi
137	36	4.2	80	6	A08907	A08907 H. sepiens (	C 210	36	4.2	3295	9	AF070673	AF070673 Homo sapi
138	36	4.2	140	3	DMRRNL	X08015 Drosophila	C 211	36	4.2	3508	3	AF023665	AF023665 Plasmodiu



212	36	4.2	3523	9	BC027936	Homo sapi	285	35	4.1	570	6	118355	118355 Sequence 10
213	36	4.2	3618	9	BC028036	Homo sapi	286	35	4.1	570	6	121342	121342 Sequence 10
214	36	4.2	3649	8	AB049934	Arabidops	287	35	4.1	570	6	140331	140331 Sequence 10
215	36	4.2	4108	10	BC008515	Mus muscu	288	35	4.1	570	8	AF137353	AF137353 Plasmu sat
216	36	4.2	4421	9	HSB802480		289	35	4.1	609	6	AR085690	AR085690 Sequence
217	36	4.2	5008	9	HSB805526		290	35	4.1	609	6	134392	134392 Sequence 9
218	36	4.2	5750	6	AX251465	Sequence	291	35	4.1	629	9	AK026557	AK026557 Homo sapi
219	36	4.2	5750	6	AX346910	Sequence	292	35	4.1	639	3	AF035657	AF035657 Strongylo
220	36	4.2	5829	6	AX346628	Sequence	293	35	4.1	645	8	AY070103	AY070103 Arabidops
221	36	4.2	6887	6	AX344607	Sequence	294	35	4.1	658	8	BC009561	BC009561 Homo sapi
222	36	4.2	8392	6	AX346392	Sequence	295	35	4.1	760	6	AR051976	AR051976 Sequence
223	36	4.2	8606	3	TPY238858	Sequence	296	35	4.1	760	6	152150	152150 Sequence 49
224	36	4.2	9731	6	AX345892	Sequence	297	35	4.1	767	3	AT118824	AT118824 Drosophill
225	36	4.2	12029	3	AE001405	Plasmodiu	298	35	4.1	791	8	OSA004966	OSA004966 Oryza sat
226	36	4.2	127046	3	AC115581	Dicystoste	299	35	4.1	792	8	AK026506	AK026506 Homo sapi
227	36	4.2	32070	2	AC116983	Dicystoste	300	35	4.1	822	8	AF395880	AF395880 Oryza sat
228	36	4.2	39369	2	AC115681	Dicystoste	301	35	4.1	843	3	AY118774	AY118774 Drosophill
229	36	4.2	43993	2	AC116965	Dicystoste	302	35	4.1	844	9	BC018536	BC018536 Homo sapi
230	36	4.2	52514	2	AC116551	Dicystoste	303	35	4.1	887	9	BC012470	BC012470 Homo sapi
231	36	4.2	53440	2	AC117078	Dicystoste	304	35	4.1	944	3	AY118333	AY118333 Drosophill
232	36	4.2	56099	2	AC115598	Dicystoste	305	35	4.1	945	3	AF457546	AF457546 Anopheles
233	36	4.2	105470	2	AC116306	Dicystoste	306	35	4.1	964	3	BC005114	BC005114 Homo sapi
234	36	4.2	106434	3	AC117080	Dicystoste	307	35	4.1	973	10	AF020524	AF020524 Mus muscu
235	36	4.2	110000	2	AC109772	Rattus no	308	35	4.1	978	9	HSB800758	HSB800758 Homo sapi
236	36	4.2	110000	2	PFMAL4P1_1	Continuaction (2 of	309	35	4.1	980	5	AF157110	AF157110 Datio rer
237	36	4.2	113880	2	PFMAL3P4	Plasmodiu	310	35	4.1	986	3	AY113296	AY113296 Drosophill
238	36	4.2	124820	2	AC117073	Dicystoste	311	35	4.1	1116	8	OSA237661	OSA237661 Oryza sat
239	36	4.2	129360	2	AC117079	Dicystoste	312	35	4.1	1160	6	AR112636	AR112636 Sequence
240	36	4.2	131397	2	CNS08C9D	AL732638 Oryza sat	313	35	4.1	1186	9	HS4420453	HS4420453 Homo sapi
241	36	4.2	132000	2	AC116976	Dicystoste	314	35	4.1	1196	8	AY130289	AY130289 Homo sapi
242	36	4.2	132000	2	AC116976	Dicystoste	315	35	4.1	1234	3	AY119240	AY119240 Arabidops
243	36	4.2	135581	2	AP000771	Homo sapi	316	35	4.1	1240	3	TBU72205	TBU72205 Arabidops
244	36	4.2	136755	2	CNS08C86	AL845344 Oryza sat	317	35	4.1	1281	10	BC029254	BC029254 Trypanosoma
245	36	4.2	144825	2	CNS08C4V	AL844498 Oryza sat	318	35	4.1	1312	6	AX306543	AX306543 Mus muscu
246	36	4.2	148312	2	AC101906	Mus muscu	319	35	4.1	1316	9	AK027106	AK027106 Homo sapi
247	36	4.2	149583	2	AC128560	Rattus no	320	35	4.1	1327	6	AX094799	AX094799 Sequence
248	36	4.2	160759	2	AC117082	Dicystoste	321	35	4.1	1334	9	BC008673	BC008673 Homo sapi
249	36	4.2	164520	2	AC020738	Homo sapi	322	35	4.1	1337	9	AF098798	AF098798 Homo sapi
250	36	4.2	165353	9	AC023822	Homo sapi	323	35	4.1	1341	9	AF029689	AF029689 Homo sapi
251	36	4.2	167303	9	AP001877	Homo sapi	324	35	4.1	1358	6	AX347225	AX347225 Sequence
252	36	4.2	170168	2	AC115313	Rattus no	325	35	4.1	1388	9	AK027096	AK027096 Homo sapi
253	36	4.2	174900	2	AC110319	Rattus no	326	35	4.1	1391	9	BC032696	BC032696 Homo sapi
254	36	4.2	178273	2	AC005308	Plasmodiu	327	35	4.1	1425	8	SOFSK93	SOFSK93 Glycine max
255	36	4.2	179335	2	AP000870	Homo sapi	328	35	4.1	1432	8	AY069596	AY069596 Drosophill
256	36	4.2	181200	3	AC099365	Rattus no	329	35	4.1	1481	3	AY061988	AY061988 Epiphyas
257	36	4.2	181635	3	AC0088224	Drosophill	330	35	4.1	1482	9	BC004991	BC004991 Homo sapi
258	36	4.2	183457	2	AC113038	Mus muscu	331	35	4.1	1495	5	AF306736	AF306736 Xenopus l
259	36	4.2	184703	2	AL773566	Mus muscu	332	35	4.1	1544	3	AF307529	AF307529 Drosophill
260	36	4.2	189303	2	AC107114	Rattus no	333	35	4.1	1550	9	HSKMR9	HSKMR9 H.sapiens m
261	36	4.2	199308	2	AC073679	Sequence	334	35	4.1	1584	10	HAJB81M	HAJB81M H. annuus Ub
262	36	4.2	200776	2	AC107841	Mus muscu	335	35	4.1	1584	10	AF083032	AF083032 Mus muscu
263	36	4.2	202785	2	AC109204	Mus muscu	336	35	4.1	1590	8	BC003614	BC003614 Homo sapi
264	36	4.2	203437	2	AC125468	Mus muscu	337	35	4.1	1613	9	AF014801	AF014801 Eschschol
265	36	4.2	212368	2	AC102048	Mus muscu	338	35	4.1	1619	9	BC032001	BC032001 Homo sapi
266	36	4.2	212139	2	AC022129	Homo sapi	339	35	4.1	1627	9	BC023977	BC023977 Homo sapi
267	36	4.2	222045	2	AC023354	Mus muscu	340	35	4.1	1629	9	BC026015	BC026015 Homo sapi
268	36	4.2	225408	2	PFMAL4P4	Sequence	341	35	4.1	1636	3	CEU60058	CEU60058 Caenorhabet
269	36	4.2	226599	2	AC079490	Mus muscu	342	35	4.1	1667	9	HSU90912	HSU90912 Human clone
270	36	4.2	229487	2	AC095711	Rattus no	343	35	4.1	1682	9	BC016047	BC016047 Homo sapi
271	36	4.2	229487	2	AC127293	Mus muscu	344	35	4.1	1682	8	AF230740	AF230740 Euphorbia
272	36	4.2	234267	2	AC127293	Mus muscu	345	35	4.1	1703	9	BC002370	BC002370 Homo sapi
273	36	4.2	233665	2	AC116966	Dicystoste	346	35	4.1	1722	9	AB048919	AB048919 Macaca fa
274	36	4.2	268147	2	AC116966	Dicystoste	347	35	4.1	1739	9	BC031645	BC031645 Homo sapi
275	36	4.2	268147	2	AC116966	Dicystoste	348	35	4.1	1740	6	A37880	A37880 Sequence 2
276	35	4.1	55	6	129928	Sequence	349	35	4.1	1740	6	A95149	A95149 Sequence 2
277	35	4.1	278	6	AX185578	Sequence	350	35	4.1	1740	6	AR157917	AR157917 Sequence
278	35	4.1	285	9	BC011247	Homo sapi	351	35	4.1	1740	6	188788	188788 Sequence 3
279	35	4.1	314	6	AX331680	Sequence	352	35	4.1	1740	8	ATAAP2	ATAAP2 X71787 A.Challiana
280	35	4.1	444	11	G21431	human STS W	353	35	4.1	1743	3	AY070515	AY070515 Drosophill
281	35	4.1	500	8	AFR105432	Sequence	354	35	4.1	1749	8	AF249299	AF249299 Cunningham
282	35	4.1	546	10	BC026533	Mus muscu	355	35	4.1	1765	10	BC034838	BC034838 Mus muscu
283	35	4.1	548	9	BC001700	Homo sapi	356	35	4.1	1784	10	BC030456	BC030456 Mus muscu
284	35	4.1	556	9	AB055361	Macaca fa	357	35	4.1	1813	9	BC011187	BC011187 Homo sapi

358	35	4.1	1833	9	BC008640	Homo sapi	431	35	4.1	2930	9	AF195092	AF195092 Homo sapi
359	35	4.1	1845	3	AY069470	Drosophila	432	35	4.1	2975	9	HSN804650	ALB33337 Homo sapi
360	35	4.1	1864	3	AB056421	Macaca fa	433	35	4.1	2992	3	AF012945	AF012945 Dictyoste
361	35	4.1	1884	3	AB088425	Bombyx mo	434	35	4.1	3037	9	HSN803746	ALB32438 Homo sapi
362	35	4.1	1892	10	AK025387	Homo sapi	435	35	4.1	3104	6	AB4105	AB4105 Sequence 9
363	35	4.1	1899	9	BC018278	Mus muscu	436	35	4.1	3152	8	BC014305	BC014305 Homo sapi
364	35	4.1	1907	9	BC033684	Homo sapi	437	35	4.1	3214	8	AF203700	AF203700 Phaseolus
365	35	4.1	1907	9	BC033684	Homo sapi	438	35	4.1	3245	9	BC013579	BC013579 Homo sapi
366	35	4.1	1915	6	AF110640	Homo sapi	439	35	4.1	3353	9	HSN803622	ALB33313 Homo sapi
367	35	4.1	1931	6	AF183234	Sequence	440	35	4.1	3379	9	HSN804444	ALB33313 Homo sapi
368	35	4.1	1931	6	AF205634	Sequence	441	35	4.1	3405	3	AY119601	AY119601 Drosophila
369	35	4.1	1950	3	AX082199	Sequence	442	35	4.1	3669	9	HSN805593	ALB3406 Homo sapi
370	35	4.1	1951	3	BC016285	Bacillus	443	35	4.1	3781	9	BC009989	BC009989 Homo sapi
371	35	4.1	1961	10	BC021360	Mus muscu	444	35	4.1	3905	6	A03736	A03736 H sapiens m
372	35	4.1	1975	3	AB072307	Bombyx mo	445	35	4.1	3910	10	BC029763	BC029763 Mus muscu
373	35	4.1	1980	3	BC033790	Homo sapi	446	35	4.1	3930	5	EC027963	EC027963 Homo sapi
374	35	4.1	1999	10	MMU80780	Mus muscu	447	35	4.1	3980	5	AF064799	AF064799 Dario rer
375	35	4.1	2005	9	BC027939	Sequence	448	35	4.1	3993	3	AF362370	AF362370 Dictyoste
376	35	4.1	2007	6	AF170139	Sequence	449	35	4.1	4018	9	HSN802205	ALB37480 Homo sapi
377	35	4.1	2007	6	AF170140	Sequence	450	35	4.1	4028	3	AF003106	AF003106 Drosophila
378	35	4.1	2029	3	AB069422	Drosophila	451	35	4.1	4045	6	AX356495	AX356495 Sequence
379	35	4.1	2048	9	BC013002	Homo sapi	452	35	4.1	4120	8	OSG002	Y12595 O sativa mR
380	35	4.1	2059	10	BC014802	Mus muscu	453	35	4.1	4193	9	AF017789	AF017789 Homo sapi
381	35	4.1	2066	9	BC002611	Sequence	454	35	4.1	4407	6	AX086676	AX086676 Sequence
382	35	4.1	2126	10	BC024858	Mus muscu	455	35	4.1	4434	9	HSN801973	ALB36805 Homo sapi
383	35	4.1	2132	10	BC026639	Sequence	456	35	4.1	4452	10	AF054831	AF054831 Mus muscu
384	35	4.1	2150	9	AB070072	Macaca fa	457	35	4.1	4574	9	HSN804034	ALB32223 Homo sapi
385	35	4.1	2163	3	AF255674	Drosophila	458	35	4.1	4629	9	HSN804034	ALB32223 Homo sapi
386	35	4.1	2207	3	AB065434	Homo sapi	459	35	4.1	4702	10	BC030304	BC030304 Mus muscu
387	35	4.1	2228	10	BC015289	Mus muscu	460	35	4.1	4759	3	DMUSHAP	Y12322 D. melanogast
388	35	4.1	2254	10	BC006652	Mus muscu	461	35	4.1	5001	3	PPAABRA	M2435 Homo sapien
389	35	4.1	2264	9	BC003065	Homo sapi	462	35	4.1	5312	3	AF241823	AF241823 Dictyoste
390	35	4.1	2279	10	BC025490	Mus muscu	463	35	4.1	5890	3	ADU60086	AY119181 Drosophila
391	35	4.1	2366	9	BC034043	Sequence 1	464	35	4.1	6045	6	AX344536	AX344536 Sequence
392	35	4.1	2400	6	AS5845	Sequence 1	465	35	4.1	6814	6	HSN803878	AX349055 Sequence
393	35	4.1	2400	8	NPZEXANT	Sequence 1	466	35	4.1	7622	3	AF112367	AX346821 Homo sapi
394	35	4.1	2403	9	AF255309	Sequence 1	467	35	4.1	8093	6	AX251189	AX112367 Plasmodu
395	35	4.1	2413	3	HSN801715	Sequence 1	468	35	4.1	8093	6	AX281166	AX251189 Sequence
396	35	4.1	2413	3	AY118685	Drosophila	469	35	4.1	8093	6	AX347410	AX281166 Sequence
397	35	4.1	2452	3	AF337815	Drosophila	470	35	4.1	8144	6	AX356411	AX349131 Sequence
398	35	4.1	2462	3	BC016388	Drosophila	471	35	4.1	8548	3	AX251388	AX349131 Sequence
399	35	4.1	2485	9	AK094962	Homo sapi	472	35	4.1	8649	6	AX251083	AX356411 Sequence
400	35	4.1	2517	9	BC011872	Homo sapi	473	35	4.1	8649	6	AX251856	AF151111 Dictyoste
401	35	4.1	2518	10	BC018260	Mus muscu	474	35	4.1	8649	6	AX344242	AX251856 Sequence
402	35	4.1	2530	9	AK000309	Sequence 1	475	35	4.1	8964	6	AX348633	AX344242 Sequence
403	35	4.1	2536	9	BC007315	Homo sapi	476	35	4.1	8964	6	AX251977	AX348633 Sequence
404	35	4.1	2546	9	AB055293	Macaca fa	477	35	4.1	8964	6	AX346337	AX251977 Sequence
405	35	4.1	2558	9	BC030525	Homo sapi	478	35	4.1	8964	6	AX346337	AX346337 Sequence
406	35	4.1	2592	9	BC012807	Homo sapi	479	35	4.1	10091	3	AC115589	AX346337 Sequence
407	35	4.1	2600	3	AY060298	Drosophila	480	35	4.1	10957	6	AX254014	AX346337 Sequence
408	35	4.1	2640	9	AF184965	Homo sapi	481	35	4.1	11523	6	AX254014	AX346337 Sequence
409	35	4.1	2649	10	HSU09366	Human zinc	482	35	4.1	11523	6	AX254014	AX346337 Sequence
410	35	4.1	2653	3	AY089530	Drosophila	483	35	4.1	11523	6	AX254014	AX346337 Sequence
411	35	4.1	2663	3	AF337552	Drosophila	484	35	4.1	11523	6	AX254014	AX346337 Sequence
412	35	4.1	2738	9	BC029891	Homo sapi	485	35	4.1	11523	6	AX254014	AX346337 Sequence
413	35	4.1	2768	5	AB044143	Cynops py	486	35	4.1	11523	6	AX254014	AX346337 Sequence
414	35	4.1	2782	9	HSN803706	Homo sapi	487	35	4.1	11523	6	AX254014	AX346337 Sequence
415	35	4.1	2807	9	AY038999	Homo sapi	488	35	4.1	11523	6	AX254014	AX346337 Sequence
416	35	4.1	2814	6	AX410773	Sequence	489	35	4.1	11523	6	AX254014	AX346337 Sequence
417	35	4.1	2814	6	AX454750	Sequence	490	35	4.1	11523	6	AX254014	AX346337 Sequence
418	35	4.1	2814	6	AX491228	Sequence	491	35	4.1	11523	6	AX254014	AX346337 Sequence
419	35	4.1	2814	6	HSU73682	Human menin	492	35	4.1	11523	6	AX254014	AX346337 Sequence
420	35	4.1	2814	6	AX410773	Sequence	493	35	4.1	11523	6	AX254014	AX346337 Sequence
421	35	4.1	2814	6	AX454750	Sequence	494	35	4.1	11523	6	AX254014	AX346337 Sequence
422	35	4.1	2814	6	AX491228	Sequence	495	35	4.1	11523	6	AX254014	AX346337 Sequence
423	35	4.1	2814	6	HSU73682	Human menin	496	35	4.1	11523	6	AX254014	AX346337 Sequence
424	35	4.1	2814	6	AX410773	Sequence	497	35	4.1	11523	6	AX254014	AX346337 Sequence
425	35	4.1	2814	6	AX454750	Sequence	498	35	4.1	11523	6	AX254014	AX346337 Sequence
426	35	4.1	2814	6	AX491228	Sequence	499	35	4.1	11523	6	AX254014	AX346337 Sequence
427	35	4.1	2814	6	HSU73682	Human menin	500	35	4.1	11523	6	AX254014	AX346337 Sequence
428	35	4.1	2814	6	AX410773	Sequence	501	35	4.1	11523	6	AX254014	AX346337 Sequence
429	35	4.1	2814	6	AX454750	Sequence	502	35	4.1	11523	6	AX254014	AX346337 Sequence
430	35	4.1	2814	6	AX491228	Sequence	503	35	4.1	11523	6	AX254014	AX346337 Sequence

504	35	4.1	14529	3	AE001397	AE001397 Plasmodiu	C 577	35	4.1	154583	10	AL672156	AL672156 Mouse DNA
505	35	4.1	14622	3	AE001421	AE001421 Plasmodiu	C 578	35	4.1	155359	9	AP003689	AP003689 Homo sapi
506	35	4.1	21537	6	AX346901	AX346901 Sequence	C 579	35	4.1	156060	2	AC004153	AC004153 Plasmodiu
507	35	4.1	27291	2	AC115575	AC115575 Dictyoste	C 580	35	4.1	156060	2	AC004153	AC004153 Plasmodiu
508	35	4.1	30638	2	AC020149	AC020149 Drosophi1	C 581	35	4.1	156121	2	CNS088CAN	ALB31810 Oryza sat
509	35	4.1	33311	9	AC004235	AC004235 Homo sapi	C 582	35	4.1	156442	9	AC107385	AC107385 Homo sapi
510	35	4.1	35276	2	AC115611	AC115611 Dictyoste	C 583	35	4.1	156461	2	AC090555	AC090555 Homo sapi
511	35	4.1	39369	2	AC115681	AC115681 Dictyoste	C 584	35	4.1	156533	2	AC117070	AC117070 Dictyoste
512	35	4.1	40770	2	AC122738	AC122738 Mus muscu	C 585	35	4.1	157083	2	AC117758	AC117758 Mus muscu
513	35	4.1	41857	2	AC126371	AC126371 Homo sapi	C 586	35	4.1	158548	3	PFMAL13P2	AL034558 Plasmodiu
514	35	4.1	45296	3	AC115613	AC115613 Dictyoste	C 587	35	4.1	158590	2	AC130197	AC130197 Felis cat
515	35	4.1	47573	2	AF030694	AF030694 Plasmodiu	C 588	35	4.1	158901	9	AC093852	AC093852 Homo sapi
516	35	4.1	47791	2	AC115593	AC115593 Dictyoste	C 589	35	4.1	159202	2	AC087347	AC087347 Homo sapi
517	35	4.1	47791	2	AC115593	AC115593 Dictyoste	C 590	35	4.1	159202	2	AC117082	AC117082 Dictyoste
518	35	4.1	53150	2	AC116305	AC116305 Dictyoste	C 591	35	4.1	160759	2	AC094329	AC094329 Dictyoste
519	35	4.1	56956	2	AC116984	AC116984 Dictyoste	C 592	35	4.1	160893	2	AC094329	AC094329 Dictyoste
520	35	4.1	58641	2	AC117081	AC117081 Dictyoste	C 593	35	4.1	162143	2	AC097553	AC097553 Dictyoste
521	35	4.1	58641	2	AC104030	AC104030 Homo sapi	C 594	35	4.1	162367	2	AC103241	AC103241 Dictyoste
522	35	4.1	59004	2	AC104030	AC104030 Homo sapi	C 595	35	4.1	162643	2	AC113327	AC113327 Mus muscu
523	35	4.1	59507	2	AC115680	AC115680 Dictyoste	C 596	35	4.1	163166	9	AL556632	AL556632 Human DNA
524	35	4.1	60182	2	AC122581	AC122581 Rattus no	C 597	35	4.1	163332	2	AC115239	AC115239 Homo sapi
525	35	4.1	60583	8	PFMAL1P6	PFMAL1P6 Plasmodiu	C 600	35	4.1	164521	2	AC119354	AC119354 Rattus no
526	35	4.1	61020	6	AB011475	AB011475 Arabidops	C 601	35	4.1	164744	2	AC018914	AC018914 Homo sapi
527	35	4.1	62352	2	AX251546	AX251546 Sequence	C 602	35	4.1	165269	2	AC117064	AC117064 Rattus no
528	35	4.1	62352	2	AC116990	AC116990 Dictyoste	C 603	35	4.1	165804	2	AC016840	AC016840 Homo sapi
529	35	4.1	62352	2	AC116990	AC116990 Dictyoste	C 604	35	4.1	165814	2	AC102296	AC102296 Mus muscu
530	35	4.1	68080	2	AC116955	AC116955 Dictyoste	C 605	35	4.1	165814	2	AC073382	AC073382 Homo sapi
531	35	4.1	69027	9	AL596306	AL596306 Human DNA	C 606	35	4.1	166333	2	AC013749	AC013749 Homo sapi
532	35	4.1	69868	2	AC102302	AC102302 Mus muscu	C 607	35	4.1	166617	10	AF332859S2	AF332859S2 Mus muscu
533	35	4.1	71580	2	AC118250	AC118250 Mus muscu	C 608	35	4.1	166676	2	AC106845	AC106845 Homo sapi
534	35	4.1	74284	2	AC034277	AC034277 Homo sapi	C 609	35	4.1	169172	2	AC026663	AC026663 Homo sapi
535	35	4.1	74850	2	AC094412	AC094412 Rattus no	C 610	35	4.1	170844	2	AL805938	AL805938 Mus muscu
536	35	4.1	77096	2	AC116030	AC116030 Dictyoste	C 611	35	4.1	171187	2	AC116960	AC116960 Dictyoste
537	35	4.1	81179	2	AC116100	AC116100 Dictyoste	C 612	35	4.1	171681	2	AC091541	AC091541 Canis fam
538	35	4.1	81179	2	AC116100	AC116100 Dictyoste	C 613	35	4.1	173920	2	AC020747	AC020747 Rattus no
539	35	4.1	93491	2	AL591388	AL591388 Dictyoste	C 614	35	4.1	173977	2	AL845165	AL845165 Mus muscu
540	35	4.1	96863	2	AL591388	AL591388 Dictyoste	C 615	35	4.1	174087	2	CNS08CB7	AL845345 Oryza sat
541	35	4.1	96870	9	AC007200	AC007200 Homo sapi	C 616	35	4.1	174087	2	AC125096	AC125096 Mus muscu
542	35	4.1	100726	2	AC116961	AC116961 Dictyoste	C 617	35	4.1	174418	2	AC098449	AC098449 Rattus no
543	35	4.1	104014	2	AC116921	AC116921 Dictyoste	C 618	35	4.1	174851	2	AL844198	AL844198 Mus muscu
544	35	4.1	105470	2	AC116306	AC116306 Dictyoste	C 619	35	4.1	176113	2	AC115133	AC115133 Rattus no
545	35	4.1	106958	9	AL807813	AL807813 Human DNA	C 620	35	4.1	177130	2	AC019351	AC019351 Homo sapi
546	35	4.1	107289	2	AC116923	AC116923 Dictyoste	C 621	35	4.1	180169	2	AC098961	AC098961 Rattus no
547	35	4.1	107686	2	AC105493	AC105493 Rattus no	C 622	35	4.1	180355	2	AC022167	AC022167 Homo sapi
548	35	4.1	107733	2	AC116979	AC116979 Dictyoste	C 623	35	4.1	180416	2	AC096793	AC096793 Rattus no
549	35	4.1	108908	3	PFMAL3P8	PFMAL3P8 Plasmodiu	C 624	35	4.1	180731	2	AC103182	AC103182 Rattus no
550	35	4.1	109342	2	AC095081	AC095081 Rattus no	C 625	35	4.1	181242	2	AL807820	AL807820 Mus muscu
551	35	4.1	110000	2	AC064053_0	AC064053 Mus muscu	C 626	35	4.1	181253	2	AC098537	AC098537 Rattus no
552	35	4.1	110000	2	AC063241_1	AC063241 Mus muscu	C 627	35	4.1	181335	2	AC016131	AC016131 Drosophi1
553	35	4.1	111187	2	AC009283	AC009283 Homo sapi	C 628	35	4.1	181383	10	AC104328	AC104328 Homo sapi
554	35	4.1	111492	2	AC103477	AC103477 Rattus no	C 629	35	4.1	183328	2	AC101773	AC101773 Mus muscu
555	35	4.1	111744	2	AC096968	AC096968 Homo sapi	C 630	35	4.1	183800	2	AL645522	AL645522 Mouse DNA
556	35	4.1	115981	2	AL591072	AL591072 Homo sapi	C 631	35	4.1	183800	2	AC123118	AC123118 Rattus no
557	35	4.1	116688	9	HSJ393E18	HSJ393E18 Human DNA	C 632	35	4.1	183800	2	AL671400	AL671400 Mouse DNA
558	35	4.1	117314	2	AC121242	AC121242 Medicago	C 633	35	4.1	184865	2	AL807824	AL807824 Mus muscu
559	35	4.1	119295	2	AC127427	AC127427 Magnapor	C 634	35	4.1	184909	2	AL844863	AL844863 Mus muscu
560	35	4.1	123820	2	AC117076	AC117076 Dictyoste	C 635	35	4.1	185165	2	AC094995	AC094995 Rattus no
561	35	4.1	123830	2	AC100244	AC100244 Mus muscu	C 636	35	4.1	186465	9	AC105052	AC105052 Homo sapi
562	35	4.1	124820	2	AC117073	AC117073 Dictyoste	C 637	35	4.1	187083	10	AC117233	AC117233 Homo sapi
563	35	4.1	127211	2	AC117066	AC117066 Rattus no	C 638	35	4.1	188391	10	AL671400	AL671400 Mouse DNA
564	35	4.1	128941	2	AC110348	AC110348 Rattus no	C 639	35	4.1	188391	2	AC109656	AC109656 Rattus no
565	35	4.1	135925	2	AC128140	AC128140 Rattus no	C 640	35	4.1	188993	2	AC093998	AC093998 Rattus no
566	35	4.1	136607	9	AC083809	AC083809 Homo sapi	C 641	35	4.1	189123	2	AC023383	AC023383 Homo sapi
567	35	4.1	137410	2	AC113861	AC113861 Rattus no	C 642	35	4.1	190225	9	AC018781	AC018781 Homo sapi
568	35	4.1	140576	2	AC120524	AC120524 Felis cat	C 643	35	4.1	190225	2	AC025283	AC025283 Homo sapi
569	35	4.1	140855	2	AC094340	AC094340 Rattus no	C 644	35	4.1	192309	2	AC115642	AC115642 Rattus no
570	35	4.1	140855	2	AC094340	AC094340 Rattus no	C 645	35	4.1	192929	2	AC005505	AC005505 Plasmodiu
571	35	4.1	141358	2	AC097196	AC097196 Rattus no	C 646	35	4.1	193149	2	AC128845	AC128845 Rattus no
572	35	4.1	146255	2	AC084170	AC084170 Homo sapi	C 647	35	4.1	193383	2	AC123256	AC123256 Rattus no
573	35	4.1	149253	2	AC012245	AC012245 Homo sapi	C 648	35	4.1	194907	2	AC106835	AC106835 Mus muscu
574	35	4.1	150060	2	AC108973	AC108973 Rattus no	C 649	35	4.1	195411	10	AL669958	AL669958 Mouse DNA
575	35	4.1	150759	2	AP001130	AP001130 Homo sapi							
576	35	4.1	151736	2	AC024439	AC024439 Homo sapi							
		4.1	152506	2	AC100830	AC100830 Homo sapi							

C 650	35	4.1 198273	2	AC107366	AC107366 Mus muscu	C 723	34	3.9	130	6	AX256038	AX256038 Sequence
C 651	35	4.1 199551	2	AC006281	AC006281 Plasmodiu	724	34	3.9	161	6	A08916	A08916 H. sapiens (
C 652	35	4.1 199960	2	AC108944	AC108944 Mus muscu	725	34	3.9	210	4	AF250317	AF250317 Canis fam
C 653	35	4.1 200789	2	AC106109	AC106109 Rattus no	C 726	34	3.9	254	6	AX334011	AX334011 Sequence
C 654	35	4.1 201724	2	AC119506	AC119506 Rattus no	C 727	34	3.9	254	6	AX337543	AX337543 Sequence
C 655	35	4.1 202374	10	AC122868	AC122868 Mus muscu	728	34	3.9	289	6	AR068753	AR068753 Sequence
C 656	35	4.1 202727	2	AC109724	AC109724 Rattus no	729	34	3.9	289	6	163077	Sequence 3
C 657	35	4.1 202740	10	AC098890	AC098890 Mus muscu	730	34	3.9	291	6	AX191241	AX191241 Sequence
C 658	35	4.1 202907	2	AC123073	AC123073 Mus muscu	731	34	3.9	291	6	AX395977	AX395977 Sequence
C 659	35	4.1 203519	10	AC112258	AC112258 Mus muscu	732	34	3.9	301	6	AF417164	AF417164 Homo sapi
C 660	35	4.1 204096	2	AC027704	AC027704 Homo sapi	733	34	3.9	317	9	AK000656	AK000656 Homo sapi
C 661	35	4.1 204652	2	PFMAL1396	PFMAL1396 Plasmodiu	734	34	3.9	324	3	AF483705	AF483705 Ixodes sc
C 662	35	4.1 205823	2	AC111087	AC111087 Mus muscu	C 735	34	3.9	327	6	AX182090	AX182090 Sequence
C 663	35	4.1 206356	2	AC112085	AC112085 Rattus no	736	34	3.9	332	6	AX186910	AX186910 Sequence
C 664	35	4.1 207392	2	AC111838	AC111838 Rattus no	C 737	34	3.9	370	6	AX046458	AX046458 Sequence
C 665	35	4.1 207456	2	AC113025	AC113025 Mus muscu	738	34	3.9	374	6	AX198829	AX198829 Sequence
C 666	35	4.1 207957	9	AC004470	AC004470 Homo sapi	739	34	3.9	374	6	AX209356	AX209356 Sequence
C 667	35	4.1 208160	2	AC122215	AC122215 Mus muscu	740	34	3.9	381	3	AF159974	AF159974 Butus ma
C 668	35	4.1 208568	2	AC110349	AC110349 Rattus no	741	34	3.9	382	9	AF207859	AF207859 Homo sapi
C 669	35	4.1 210423	2	AL672074	AL672074 Mus muscu	742	34	3.9	384	3	AF483734	AF483734 Ixodes sc
C 670	35	4.1 213171	2	AC124200	AC124200 Mus muscu	743	34	3.9	385	3	AF151796	AF151796 Butus ma
C 671	35	4.1 213251	2	AC116458	AC116458 Mus muscu	744	34	3.9	386	3	AF495660	AF495660 Varetia v
C 672	35	4.1 213303	10	AL671520	AL671520 Mouse DNA	745	34	3.9	409	3	AF156597	AF156597 Butus ma
C 673	35	4.1 215115	2	AC113311	AC113311 Mus muscu	746	34	3.9	426	9	HSM802848	HSM802848 Sequence
C 674	35	4.1 216920	2	AC094599	AC094599 Rattus no	747	34	3.9	427	6	AX070822	AX070822 Drosophila
C 675	35	4.1 216937	2	AC022061	AC022061 Mus muscu	C 748	34	3.9	427	6	AX330416	AX330416 Sequence
C 676	35	4.1 217014	2	AC104549	AC104549 Mus muscu	749	34	3.9	431	3	AF483662	AF483662 Ixodes sc
C 677	35	4.1 219519	2	AC115753	AC115753 Mus muscu	C 750	34	3.9	446	3	AF146743	AF146743 Mesobuthu
C 678	35	4.1 221548	2	AC099714	AC099714 Mus muscu	C 751	34	3.9	446	11	G21964	G21964 human SRS W
C 679	35	4.1 222618	2	AC110335	AC110335 Rattus no	C 752	34	3.9	452	11	G24486	G24486 human SRS W
C 680	35	4.1 224149	2	AC119799	AC119799 Pan trogl	C 753	34	3.9	457	11	G24145	G24145 human SRS W
C 681	35	4.1 225204	2	AL772138	AL772138 Mus muscu	754	34	3.9	458	6	AR077207	AR077207 Sequence
C 682	35	4.1 227617	2	AC125352	AC125352 Mus muscu	755	34	3.9	458	6	AR105068	AR105068 Sequence
C 683	35	4.1 228023	2	AL672184	AL672184 Mus muscu	C 756	34	3.9	458	11	G27046	G27046 human SRS S
C 684	35	4.1 232793	2	AC107719	AC107719 Mus muscu	C 757	34	3.9	460	3	AY113221	AY113221 Drosophila
C 685	35	4.1 232793	2	AC107719	AC107719 Mus muscu	C 758	34	3.9	463	8	AF333788	AF333788 Pirus tae
C 686	35	4.1 236029	2	AC019322	AC019322 Homo sapi	C 759	34	3.9	465	9	BC007974	BC007974 Homo sapi
C 687	35	4.1 240065	2	AC093481	AC093481 Mus muscu	C 760	34	3.9	475	6	AI4236	AI4236 D5 anonymu
C 688	35	4.1 240864	2	AC006510	AC006510 Homo sapi	761	34	3.9	475	9	BC031580	BC031580 Homo sapi
C 689	35	4.1 242218	2	AC104917	AC104917 Mus muscu	762	34	3.9	479	3	DDCOX6R	DDCOX6R Dictyosteli
C 690	35	4.1 242835	2	AC079243	AC079243 Mus muscu	763	34	3.9	490	10	AF172930	AF172930 Mus muscu
C 691	35	4.1 242979	2	AC093478	AC093478 Mus muscu	764	34	3.9	493	8	NTTIMPA	NTTIMPA N. tabacum T
C 692	35	4.1 253756	2	AC093975	AC093975 Rattus no	765	34	3.9	506	3	AF072331	AF072331 Schistos
C 693	35	4.1 256774	2	AC115964	AC115964 Dictyoste	766	34	3.9	521	10	AY028425	AY028425 Mus muscu
C 694	35	4.1 256774	2	AC115964	AC115964 Dictyoste	767	34	3.9	522	9	AF417129	AF417129 Homo sapi
C 695	35	4.1 261381	2	AL596386	AL596386 Mus muscu	768	34	3.9	530	9	BC022858	BC022858 Homo sapi
C 696	35	4.1 261381	2	AL596386	AL596386 Mus muscu	769	34	3.9	531	8	AF140500	AF140500 Oryza sat
C 697	35	4.1 263662	2	AC114344	AC114344 Rattus no	770	34	3.9	560	8	PR249397	PR249397 Phleum pr
C 698	35	4.1 263775	2	AC115418	AC115418 Rattus no	771	34	3.9	561	9	IR0687685	IR0687685 Homo sapi
C 699	35	4.1 268578	2	AC094863	AC094863 Rattus no	772	34	3.9	570	9	AF277181	AF277181 Homo sapi
C 700	35	4.1 271222	2	AC124770	AC124770 Mus muscu	773	34	3.9	586	3	AY066029	AY066029 Branchio
C 701	35	4.1 278082	2	AC111365	AC111365 Rattus no	774	34	3.9	591	3	CCMASPRRO	CCMASPRRO Ceratilis c
C 702	35	4.1 278307	2	AC118873	AC118873 Rattus no	C 775	34	3.9	592	6	AX421572	AX421572 Sequence
C 703	35	4.1 280982	2	AC1125184	AC1125184 Mus muscu	776	34	3.9	596	9	AF417145	AF417145 Homo sapi
C 704	35	4.1 287820	2	AC096329	AC096329 Rattus no	777	34	3.9	602	9	BC015519	BC015519 Homo sapi
C 705	35	4.1 291545	3	AE003600	AE003600 Drosophila	778	34	3.9	607	5	AF321772	AF321772 Ophiophag
C 706	35	4.1 293431	2	PFMAL1394	PFMAL1394 Plasmodiu	779	34	3.9	610	8	AB047923	AB047923 Oryza sat
C 707	35	4.1 308692	2	AC098559	AC098559 Rattus no	780	34	3.9	611	6	AX187051	AX187051 Sequence
C 708	35	4.1 309233	2	AC098557	AC098557 Rattus no	781	34	3.9	611	10	BC026656	BC026656 Mus muscu
C 709	35	4.1 310779	2	AC005140	AC005140 Plasmodiu	782	34	3.9	613	10	S76337	S76337 IGF-binding
C 710	35	4.1 318930	2	AC073495	AC073495 Mus muscu	783	34	3.9	613	10	BC028889	BC028889 Mus muscu
C 711	35	4.1 321003	2	PFMAL4P3	PFMAL4P3 Plasmodiu	784	34	3.9	625	8	AF110779	AF110779 Lillium lo
C 712	35	4.1 349980	6	AX344554	AX344554 Sequence	C 785	34	3.9	631	6	AX338092	AX338092 Sequence
C 713	35	4.1 349980	6	AX344554	AX344554 Sequence	786	34	3.9	633	6	AR038867	AR038867 Sequence
C 714	35	4.1 349980	6	AX344557	AX344557 Sequence	787	34	3.9	655	3	AF420265	AF420265 Anopheles
C 715	35	4.1 349980	6	AX344560	AX344560 Sequence	C 788	34	3.9	655	6	AX185692	AX185692 Sequence
C 716	35	4.1 349980	6	AX344563	AX344563 Sequence	789	34	3.9	665	3	BC021122	BC021122 Homo sapi
C 717	35	4.1 349980	6	AX344564	AX344564 Sequence	790	34	3.9	666	3	AY047580	AY047580 Drosophila
C 718	35	4.1 349980	6	AX344565	AX344565 Sequence	791	34	3.9	669	9	HSN800251	HSN800251 Homo sapi
C 719	35	4.1 349980	6	AX344569	AX344569 Sequence	792	34	3.9	673	8	CRPSP27	CRPSP27 C. reinhardt
C 720	34	3.9 77	6	A08908	A08908 H. sapiens (	793	34	3.9	677	9	AB055315	AB055315 Macaca fa
C 721	34	3.9 96	6	A08909	A08909 H. sapiens (	794	34	3.9	677	9	BC018211	BC018211 Homo sapi
C 722	34	3.9 104	6	A08910	A08910 H. sapiens (	795	34	3.9	690	9	BC021234	BC021234 Homo sapi

796	34	3.9	703	6	AX259734	Sequence	869	34	3.9	973	9	BC017067	BC017067 Homo sapi
797	34	3.9	707	3	DROCRPI	M62398 Drosophila	C 870	34	3.9	975	3	DD087514	U85184 Dictyosteli
798	34	3.9	708	8	PSGDCRG	X54377 P.sativum g	871	34	3.9	978	3	AY118474	AY118474 Drosophi
799	34	3.9	709	6	AX351026	Sequence	872	34	3.9	987	3	AY122185	AY122185 Drosophi
800	34	3.9	711	3	DVU26847	U26847 Drosophila	873	34	3.9	987	3	SCMGST	LO6180 Schistosoma
801	34	3.9	713	3	AY118450	Drosophila	874	34	3.9	989	3	AY089478	AY089478 Drosophi
802	34	3.9	714	8	AF220198	Vlt1a r1p	875	34	3.9	991	9	BC020915	U70494 Mus muscu
803	34	3.9	733	8	AF401200	Bauhinia	876	34	3.9	994	3	AY069151	BC020915 Homo sapi
804	34	3.9	734	3	AY118721	Drosophila	877	34	3.9	999	9	BC001109	AY069151 Drosophi
805	34	3.9	739	3	AY113352	Drosophila	878	34	3.9	1009	3	AF030628	BC001109 Homo sapi
806	34	3.9	746	6	AR137703	Sequence	879	34	3.9	1009	3	AF030628	AF030628 Plasmodi
807	34	3.9	748	6	A21101	N.tabacum S	880	34	3.9	1012	9	BC009540	BC009540 Homo sapi
808	34	3.9	748	6	144072	Sequence 3	881	34	3.9	1012	10	BC004052	BC004052 Mus muscu
809	34	3.9	749	3	HSA420433	AY420433 Homo sapi	C 882	34	3.9	1015	2	AB064665	AB064665 Homo sapi
810	34	3.9	758	3	AY060876	Drosophila	883	34	3.9	1016	2	AC014440	AC014440 Drosophi
811	34	3.9	759	3	AY118716	AY118716 Drosophi	884	34	3.9	1017	4	CFMALL	X92505 C.familiari
812	34	3.9	761	11	CNS06KKG	AL402998 t7 end of	885	34	3.9	1018	3	AY060602	AY060602 Drosophi
813	34	3.9	762	3	BC029884	Homo sapi	886	34	3.9	1029	3	AY113302	AY113302 Drosophi
814	34	3.9	769	8	ATU60981	U60981 Arabidopsis	887	34	3.9	1033	6	AX358283	AX358283 Sequence
815	34	3.9	770	10	BC024391	Mus muscu	888	34	3.9	1033	8	AF062915	AF062915 Arabidops
816	34	3.9	771	8	BC000608	BC000608 Homo sapi	889	34	3.9	1038	5	AY078168	AY078168 Dantio ter
817	34	3.9	779	8	PYU63631	U63631 Fragaria x	890	34	3.9	1049	9	BC011461	BC011461 Homo sapi
818	34	3.9	781	9	AF237812	Homo sapi	C 891	34	3.9	1055	4	BTCASA1	X82401 B. taurus CA
819	34	3.9	793	10	BC022621	Mus muscu	892	34	3.9	1059	10	BC025911	BC025911 Mus muscu
820	34	3.9	796	6	A23630	A23630 Corn anther	893	34	3.9	1069	9	BC009520	BC009520 Homo sapi
821	34	3.9	796	6	133111	133111 Ixodes sc	894	34	3.9	1071	9	AF339776	AF339776 Homo sapi
822	34	3.9	814	3	AF483742	AF483742 Ixodes sc	895	34	3.9	1074	3	DDICABP	M36174 D. discoideu
823	34	3.9	815	9	HSA406952	AX406952 Homo sapi	896	34	3.9	1089	8	STPSBPCEN	X99320 S. tuberosum
824	34	3.9	817	3	AF132912	Drosophila	897	34	3.9	1090	3	DV4A	X15381 Dictyosteli
825	34	3.9	818	9	BC032631	Homo sapi	898	34	3.9	1091	9	BC000745	BC000745 Homo sapi
826	34	3.9	820	10	BC033606	Mus muscu	899	34	3.9	1092	3	AY070897	AY070897 Drosophi
827	34	3.9	829	8	AF339780	Homo sapi	900	34	3.9	1094	10	AF237620	AF237620 Mus muscu
828	34	3.9	830	8	AF303940	Glycine m	901	34	3.9	1096	3	AY118489	AY118489 Drosophi
829	34	3.9	835	10	AF260436	AF260436 Rattus no	902	34	3.9	1096	10	BC032962	BC032962 Mus muscu
830	34	3.9	836	9	BC013035	BC013035 Homo sapi	903	34	3.9	1098	8	BOMIPATCP	BC032362 Homo sapi
831	34	3.9	837	8	NPSODM	X55974 N.plumbagin	904	34	3.9	1104	8	AY078282	X95639 B. oleracea
832	34	3.9	840	9	BC006414	BC006414 Homo sapi	905	34	3.9	1105	8	AY078282	AY078282 Thalassio
833	34	3.9	843	4	AF319463	AF319463 Macropus	906	34	3.9	1107	9	BC000777	BC000777 Homo sapi
834	34	3.9	845	3	AF047612	AF047612 Euroglyph	907	34	3.9	1110	9	BC006305	BC006305 Homo sapi
835	34	3.9	852	3	AY118642	AY118642 Drosophi	908	34	3.9	1111	5	AF143806	AF143806 Gallus ga
836	34	3.9	853	8	SSEPDIINH	X96361 S.pombe SPD	909	34	3.9	1111	9	BC000853	BC000853 Homo sapi
837	34	3.9	853	8	CPCRMA	Y08240 C.pteroides	910	34	3.9	1115	5	AF460212	AY119480 Drosophi
838	34	3.9	857	10	BC023139	BC023139 Mus muscu	911	34	3.9	1115	5	AF460212	AF460212 Crocodylu
839	34	3.9	864	9	HSCBICPI1	Y10776 H.sapiens m	912	34	3.9	1116	9	BC032317	BC032317 Homo sapi
840	34	3.9	871	3	AY089389	Drosophila	913	34	3.9	1118	9	BC012595	BC012595 Homo sapi
841	34	3.9	875	9	BC030260	BC030260 Homo sapi	914	34	3.9	1128	3	AY119027	AY119027 Drosophi
842	34	3.9	876	8	AF288464	Potteresi	915	34	3.9	1128	9	AF274025	AF274025 Homo sapi
843	34	3.9	879	9	BC007021	Homo sapi	916	34	3.9	1129	9	BC032768	BC032768 Homo sapi
844	34	3.9	889	9	BC016970	Homo sapi	917	34	3.9	1129	5	BC004117	BC004117 Homo sapi
845	34	3.9	899	6	AF368271	Homo sapi	918	34	3.9	1130	6	AX244574	AY118699 Drosophi
846	34	3.9	900	6	AR105095	Sequence	919	34	3.9	1131	6	AX244574	AX244574 Sequence
847	34	3.9	900	8	AF141879	Oryza sat	920	34	3.9	1133	6	AF271660	AX003386 Sequence
848	34	3.9	900	8	AF244659	AF244659 Zea mays	921	34	3.9	1133	8	AF529691	AF271660 Vlt1a bet
849	34	3.9	903	8	AY118755	Drosophila	922	34	3.9	1133	8	AF529691	AF529691 Brassica
850	34	3.9	905	3	BC015507	Homo sapi	923	34	3.9	1134	17	AF116670	AF116670 Homo sapi
851	34	3.9	907	3	AY118472	Drosophila	924	34	3.9	1136	6	E01349	E01349 CDNA encodi
852	34	3.9	910	3	BC014198	Homo sapi	925	34	3.9	1136	6	108009	108009 Sequence 3
853	34	3.9	913	3	AY118492	Drosophila	926	34	3.9	1142	3	AY089349	AY089349 Drosophi
854	34	3.9	918	8	AF017363	AF017363 Oryza sat	927	34	3.9	1143	9	BC007389	BC007389 Homo sapi
855	34	3.9	929	9	D82059	D82059 Homo sapien	928	34	3.9	1146	6	AR122651	AR122651 Sequence
856	34	3.9	930	10	BC028754	Mus muscu	929	34	3.9	1147	9	BC009016	BC009016 Homo sapi
857	34	3.9	945	9	HSW801971	AL137301 Homo sapi	930	34	3.9	1152	8	AHLHAH	X74732 A.hypochoind
858	34	3.9	948	5	AY040528	Gallus ga	931	34	3.9	1154	6	AR094308	AR094308 Sequence
859	34	3.9	951	9	AF320777	Homo sapi	932	34	3.9	1156	9	BC032114	BC032114 Homo sapi
860	34	3.9	959	6	AX052820	Sequence	933	34	3.9	1159	10	AF331040	AF331040 Mus muscu
861	34	3.9	959	9	HSMPEZE	X61970 H.sapiens m	934	34	3.9	1161	8	AF460211	AF460211 Crocodylu
862	34	3.9	961	8	NTAP24	X65700 N.tabacum m	935	34	3.9	1162	8	MNHRBCSA	M95930 Manihot esc
863	34	3.9	962	10	BC027572	Mus muscu	936	34	3.9	1168	10	BC019114	BC019114 Mus muscu
864	34	3.9	962	10	BC034133	Mus muscu	937	34	3.9	1173	10	BC014724	BC014724 Dictyostell
865	34	3.9	965	3	AY119476	Drosophila	938	34	3.9	1182	3	DDU48706	U48706 Dictyostell
866	34	3.9	965	10	AF106945	Rattus no	939	34	3.9	1182	9	HUMIZOI	M93009 Human L-iso
867	34	3.9	971	10	BC030946	Mus muscu	940	34	3.9	1190	6	AX275042	AX275042 Sequence
868	34	3.9	971	10	BC030946	Mus muscu	941	34	3.9	1191	3	AY118797	AY118797 Drosophi

## ALIGNMENTS

REMARK  
COMMENT

REMARK  
COMMENT  
Nih-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: The I.M.A.G.E. Consortium (ILNL)  
<http://www.systemsbiology.org>  
contact: [amadansystemsbiology.org](mailto:amadansystemsbiology.org)



• NIESKVS KÖLN KKYFSVERI ÖR KKD ITÖLSKHKHTVME DKVEVVPKÖPTVIDLFT

cdna@fruitfly.berkeley.edu.  
Location/Qualifiers

Location/Qualifiers











GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2003, 08:37:15 : Search time 234 Seconds  
(without alignments)  
8305.447 Million cell updates/sec

Title: US-10-021-811-35  
Perfect score: 863  
Sequence: 1 gccacgagctctacacacac.....aaaaaaaaaaaaaaaa 863

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : N\_Geneseq\_101002:\*

```
1: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDs2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	4.8	2037	24	AAD24013 Human drug metabo
2	41	4.8	2307	19	AAV44318 Mize search-associ
3	41	4.8	4329	19	AAV44319 Seq ID #7 from DB1
4	40	4.6	10825	24	ABN80207 Human chemically m
5	39	4.5	398	22	AA188050 Human polynucleoti
6	39	4.5	1061	24	ABO21136 Oligonucleotide fo
7	39	4.5	1061	24	ABO21137 Oligonucleotide fo
8	39	4.5	1402	22	AA157850 Human polynucleoti
9	39	4.5	3001	24	ABN80333 Human chemically m

C 10	39	4.5	5273	24	ABL32875 Human immune syste
C 11	39	4.5	5313	24	ABL32763 Human immune syste
C 12	39	4.5	6120	24	ABN80302 Human chemically m
C 13	39	4.5	6120	24	ABK28439 DNA transcription
C 14	39	4.5	6956	24	ABL70225 Chemically treated
C 15	39	4.5	7046	24	ABN90288 Human chemically m
C 16	39	4.5	7046	24	ABL34116 Human immune syste
C 17	39	4.5	7890	24	ABL70240 Chemically treated
C 18	39	4.5	7890	24	AA561185 Human gene regulat
C 19	39	4.5	8246	24	ABL32203 Human immune syste
C 20	39	4.5	8372	22	AA546353 Tumour suppressor
C 21	39	4.5	9963	24	ABL32694 Human immune syste
C 22	39	4.5	11178	24	ABL70507 Chemically treated
C 23	39	4.5	11178	24	AA561059 Human gene regulat
C 24	39	4.5	11178	24	ABK31172 Signal transductio
C 25	39	4.5	13427	24	ABL33926 Human immune syste
C 26	39	4.5	15500	22	AA546508 Tumour suppressor
C 27	39	4.5	17234	24	ABO67018 Human angiogenesis
C 28	39	4.5	17389	24	ABL33414 Human immune syste
C 29	38	4.4	379	22	ABL20476 Human breast cance
C 30	38	4.4	5447	22	AA546757 Tumour suppressor
C 31	38	4.4	9770	24	ABL32032 Human immune syste
C 32	38	4.4	19798	22	AAF83671 Nucleic acid seque
C 33	38	4.4	19912	22	AAF83670 Nucleic acid seque
C 34	38	4.4	20160	22	AAF83672 Nucleic acid seque
C 35	38	4.4	20217	22	AAF83674 Nucleic acid seque
C 36	38	4.4	20247	22	AAF83675 Nucleic acid seque
C 37	38	4.4	20316	22	AAF83673 Nucleic acid seque
C 38	38	4.4	27960	22	AA69779 Human immune/haema
C 39	38	4.4	27960	22	AAK73320 Human ubiqutin co
C 40	37	4.3	1137	18	AA63234 Nucleotide sequenc
C 41	37	4.3	1159	22	AAF79920 Human ubiqutin co
C 42	37	4.3	1856	18	AA63233 Human UB21 CDNA s
C 43	37	4.3	1856	24	ABO60785 Tumour necrosis fa
C 44	37	4.3	3964	19	AAV62673 Sequence encoding
C 45	37	4.3	4051	19	AAV62672 Kinechohre protei
C 46	37	4.3	5108	14	AAQ53403 Tumour suppressor
C 47	37	4.3	10136	17	AA634578 Human nervous syst
C 48	37	4.3	10136	22	AA617042 Human pancreatic r
C 49	36	4.2	152	22	AA616670 Genomic sequence #
C 50	36	4.2	152	22	AA529169 Human reproductive
C 51	36	4.2	152	22	AA606722 Human digestive sy
C 52	36	4.2	152	22	AAK89972 Human breast or ov
C 53	36	4.2	152	22	AA162628 Human prostate exp
C 54	36	4.2	152	23	ABV07534 Human prostate exp
C 55	36	4.2	237	23	ABV58043 Human prostate exp
C 56	36	4.2	257	23	ABV59128 Human cervical can
C 57	36	4.2	305	23	AAH71523 Human CDNA SRQ ID
C 58	36	4.2	332	22	ABR06391 Human polynucleoti
C 59	36	4.2	401	22	AA187701 Human prostate exp
C 60	36	4.2	491	22	AA58464 Human colon cancer
C 61	36	4.2	528	21	AA280560 CDNA clone APPLF.
C 62	36	4.2	572	21	AA172857 Human immune/haema
C 63	36	4.2	880	21	AAK61917 Flea saliv protei
C 64	36	4.2	1007	18	AA792823 Flea saliv protei
C 65	36	4.2	1007	18	AAV73412 Human colon cancer
C 66	36	4.2	1037	22	AAH34174 Human secreted pro
C 67	36	4.2	1037	21	AAAC93387 Lung cancer associ
C 68	36	4.2	1131	21	AAFI18268 Human polynucleoti
C 69	36	4.2	1230	24	AB189756 Human cancer cell
C 70	36	4.2	1254	24	AA144826 Human cancer cell
C 71	36	4.2	1384	24	AA144827 Human secreted pro
C 72	36	4.2	1384	24	AAQ20842 Mec-4 gene. Caeno
C 73	36	4.2	1542	22	AAQ20842 Mec-4 gene. Caeno
C 74	36	4.2	1606	13	AAQ38934 Murine CRM-2 codi
C 75	36	4.2	1607	21	AA595305 cDNA encoding nove
C 76	36	4.2	1626	21	AA541104 cDNA encoding nove
C 77	36	4.2	1736	20	AA586292 Complement of cDNA
C 78	36	4.2	1736	20	AA586293 Flea epoxide hydro
C 79	36	4.2	1736	22	AA166761 Flea epoxide hydro
C 80	36	4.2	1736	22	AA166762 Flea epoxide hydro
C 81	36	4.2	1736	22	AA166762 Flea epoxide hydro
C 82	36	4.2	1736	22	AAAC91978 Flea epoxide hydro

C	83	36	4.2	1736	22	AAC91979	Flea epoxide hydro
C	84	36	4.2	1736	22	AAC86292	Flea epoxide hydro
C	85	36	4.2	1736	22	AAC86293	Flea epoxide hydro
C	86	36	4.2	2070	21	AAC76167	Human ORFX ORF1722
C	87	36	4.2	2325	22	AAD08296	Human secreted pro
C	88	36	4.2	2422	21	AAC76970	Human ORFX ORF25
C	89	36	4.2	2524	21	AAC77206	Human ORFX ORF2761
C	90	36	4.2	3989	22	AAH34823	Human colon cancer
C	91	36	4.2	5750	22	AAH36708	Tumour suppressor
C	92	36	4.2	5750	24	ABL34008	Human immune syste
C	93	36	4.2	5829	24	ABL33726	Human immune syste
C	94	36	4.2	6887	24	ABN80015	Human chemically m
C	95	36	4.2	8392	24	ABL33490	Human immune syste
C	96	36	4.2	9731	24	ABL32990	Human immune syste
C	97	35	4.1	177	23	AAH36957	Human prostate exp
C	98	35	4.1	232	22	AAH33958	Human colon cancer
C	99	35	4.1	239	23	ABV56818	Human prostate exp
C	100	35	4.1	256	23	ABV52378	Human prostate exp
C	101	35	4.1	264	24	ABL87207	Human prostate exp
C	102	35	4.1	278	22	AAH69999	Human cervical can
C	103	35	4.1	301	23	ABV19472	Human prostate exp
C	104	35	4.1	307	24	ABL87642	Human ovarian can
C	105	35	4.1	314	24	ABL63852	Breast cancer rela
C	106	35	4.1	318	23	ABV56059	Human prostate exp
C	107	35	4.1	335	22	AAI19579	Human breast cance
C	108	35	4.1	341	23	ABV18579	Human prostate exp
C	109	35	4.1	345	23	ABV07703	Human prostate exp
C	110	35	4.1	348	22	AAI90986	Human polynucleoti
C	111	35	4.1	355	24	ABL81624	Human polynucleoti
C	112	35	4.1	370	23	ABV13480	Human prostate can
C	113	35	4.1	382	22	AAI93776	Human polynucleoti
C	114	35	4.1	383	22	AAI92810	Human polynucleoti
C	115	35	4.1	385	22	AAI87439	Human polynucleoti
C	116	35	4.1	386	22	AAI80166	Human polynucleoti
C	117	35	4.1	386	22	AAI92904	Human polynucleoti
C	118	35	4.1	387	23	ABV57769	Human prostate exp
C	119	35	4.1	389	22	AAI88648	Human polynucleoti
C	120	35	4.1	394	23	ABV57930	Human prostate exp
C	121	35	4.1	400	22	AAI90914	Human polynucleoti
C	122	35	4.1	401	22	AAI82944	Human polynucleoti
C	123	35	4.1	406	22	AAI87091	Human polynucleoti
C	124	35	4.1	408	22	AAI89257	Human polynucleoti
C	125	35	4.1	408	22	AAI91282	Human polynucleoti
C	126	35	4.1	409	22	AAI81446	Human polynucleoti
C	127	35	4.1	412	22	AAI87923	Human polynucleoti
C	128	35	4.1	414	22	AAI87988	Human polynucleoti
C	129	35	4.1	415	22	AAI92174	Human polynucleoti
C	130	35	4.1	430	22	AAI86288	Human polynucleoti
C	131	35	4.1	434	22	AAI82808	Human polynucleoti
C	132	35	4.1	440	23	ABV42126	Human prostate exp
C	133	35	4.1	442	23	ABV46394	Human prostate exp
C	134	35	4.1	443	23	ABV48362	Human prostate exp
C	135	35	4.1	449	23	ABV58812	Human prostate exp
C	136	35	4.1	451	22	AAK88891	Human digestive sy
C	137	35	4.1	451	22	AAI57612	Human colorectal c
C	138	35	4.1	469	23	ABV56638	Human prostate exp
C	139	35	4.1	478	22	AAI87372	Human polynucleoti
C	140	35	4.1	499	21	AAAC59599	Human secreted pro
C	141	35	4.1	499	22	AAI93507	Domestic mite B5
C	142	35	4.1	509	22	AAAD06268	Human colon cancer
C	143	35	4.1	509	24	AAAD30715	Human prostate exp
C	144	35	4.1	521	22	AAH34685	Human prostate exp
C	145	35	4.1	540	23	ABV58372	Human prostate exp
C	146	35	4.1	570	17	AAAT30258	Human prostate exp
C	147	35	4.1	570	17	AAAT3041	Cotton fibre speci
C	148	35	4.1	570	18	AAAT70048	Cotton fibre speci
C	149	35	4.1	609	18	AAAT6217	Cotton fibre speci
C	150	35	4.1	609	21	AAH23552	Cotton fibre speci
C	151	35	4.1	610	24	ABQ25734	cDNA sequence a co
C	152	35	4.1	611	24	ABQ25735	Oligonucleotide fo
C	153	35	4.1	610	23	ABV56623	Oligonucleotide fo
C	154	35	4.1	724	24	ABQ43744	Human prostate exp
C	155	35	4.1	724	24	ABQ43745	Oligonucleotide fo
C	156	35	4.1	733	21	AAC79892	Human secreted pro
C	157	35	4.1	760	16	AAQ92526	P. communis (pear)
C	158	35	4.1	762	24	ABI99925	R. mucocarpalis ce
C	159	35	4.1	767	24	ABN99160	Arabidopsis thalia
C	160	35	4.1	781	22	AAI97584	Arabidopsis thalia
C	161	35	4.1	785	24	ABN99094	Human neuroblastom
C	162	35	4.1	805	22	AAI97585	Human neuroblastom
C	163	35	4.1	811	22	AAD08501	Human secreted pro
C	164	35	4.1	839	22	AAI20429	Human secreted pro
C	165	35	4.1	863	22	AAI26117	Human breast cance
C	166	35	4.1	1051	22	AAI25968	Human cDNA encodin
C	167	35	4.1	1150	21	AAH40524	Human cDNA encodin
C	168	35	4.1	1160	19	AAV40008	Human nervous syst
C	169	35	4.1	1254	21	AAC99826	Human cell death a
C	170	35	4.1	1312	24	AAD24774	Human secreted pro
C	171	35	4.1	1327	22	AAAS02004	Glycine max ankylr
C	172	35	4.1	1358	24	ABL34323	DNA insert of a to
C	173	35	4.1	1361	21	AAH463399	Human immune syste
C	174	35	4.1	1367	20	AAH51727	Human secreted pro
C	175	35	4.1	1421	21	AAC59550	DNA encoding a hum
C	176	35	4.1	1486	20	AAI19855	Human secreted pro
C	177	35	4.1	1502	21	AAC69488	Human adult brain
C	178	35	4.1	1633	21	AAC78055	Human secreted pro
C	179	35	4.1	1675	23	AAH86356	Human cancer assoc
C	180	35	4.1	1686	20	AAV73012	DNA encoding novel
C	181	35	4.1	1686	24	ABO92037	Human adult brain
C	182	35	4.1	1763	21	AAH297043	Human polynucleoti
C	183	35	4.1	1787	21	AAH297043	Human secreted pro
C	184	35	4.1	1931	20	AAH209246	Human secreted pro
C	185	35	4.1	1931	22	AAH30001	Human CARD-3 cDNA.
C	186	35	4.1	1931	20	AAH40752	CDNA of human CARD
C	187	35	4.1	1958	21	AAH40706	Human cDNA encodin
C	188	35	4.1	2007	19	AAV40754	C. felis esterase,
C	189	35	4.1	2007	19	AAV40755	C. felis esterase,
C	190	35	4.1	2007	22	AAH21174	Ctenocephalides fe
C	191	35	4.1	2007	22	AAH21174	Human colon cancer
C	192	35	4.1	2044	22	AAH25958	Human cDNA encodin
C	193	35	4.1	2051	21	AAH25958	Rhizopus PGI1 prom
C	194	35	4.1	2138	12	AAH25958	Human secreted pro
C	195	35	4.1	2163	21	AAH25958	Human prostate can
C	196	35	4.1	2211	23	ABK43549	DNA encoding novel
C	197	35	4.1	2329	22	AAH25958	Human secreted pro
C	198	35	4.1	2355	22	AAH25958	Human secreted pro
C	199	35	4.1	2510	18	AAH25958	Nicotiana glauca
C	200	35	4.1	2512	22	AAH25958	Human colon cancer
C	201	35	4.1	2532	21	AAH25958	Human colon cancer
C	202	35	4.1	2533	21	AAH25958	Human cancer assoc
C	203	35	4.1	2814	24	ABN96922	Gene #3420 used to
C	204	35	4.1	2814	24	ABN96922	Human angiogenesis
C	205	35	4.1	2814	24	ABN96922	Ovary cancer relat
C	206	35	4.1	2814	24	ABN96922	Human PRO4527 cDNA
C	207	35	4.1	2814	24	ABN96922	DNA encoding novel
C	208	35	4.1	2863	23	ABK43677	Human h-NIMB-R enc
C	209	35	4.1	3104	9	AAH70356	Sequence encoding
C	210	35	4.1	3905	13	AAH70461	Human Factor XIII
C	211	35	4.1	3905	13	AAH70461	Human Factor XIII
C	212	35	4.1	4045	24	ABN92230	Chemically treated
C	213	35	4.1	4045	24	ABN92230	Chemically treated
C	214	35	4.1	6045	24	ABN92230	Chemically treated
C	215	35	4.1	6045	24	ABN92230	Chemically treated
C	216	35	4.1	7204	24	ABN92230	Signal transductio
C	217	35	4.1	8093	24	ABN92230	Human immune syste
C	218	35	4.1	8093	24	ABN92230	Human immune syste
C	219	35	4.1	8093	24	ABN92230	Human immune syste
C	220	35	4.1	8093	24	ABN92230	Human immune syste
C	221	35	4.1	8144	22	AAH46634	Chemically treated
C	222	35	4.1	8649	22	AAH46634	Human DNA for stag
C	223	35	4.1	8649	22	AAH46634	Human MHL gene as
C	224	35	4.1	8649	22	AAH46634	Tumour suppressor
C	225	35	4.1	8649	24	AAH46634	Tumour suppressor
C	226	35	4.1	8649	24	AAH46634	Chemically treated
C	227	35	4.1	8964	24	ABN92230	Human gene regulat
C	228	35	4.1	8964	24	ABN92230	Human gene regulat

C 229	35	4.1	8964	24	ABK31373	Signal transductio	C 302	34	3.9	323	22	AAI23920	Human breast cance
230	35	4.1	10957	22	ABL33110	Human immune syste	C 303	34	3.9	327	22	AAI07726	Cervical cancer pr
C 231	35	4.1	11523	24	AA545467	Chemically pretrea	C 304	34	3.9	332	22	AAH71330	Human cervical can
C 232	35	4.1	11523	22	AA546641	Tumour suppressor	305	34	3.9	337	23	ABV58905	Human prostate exp
C 233	35	4.1	11523	24	ABK34001	Human DNA for steg	306	34	3.9	337	23	ABV54338	Human prostate exp
C 234	35	4.1	11523	24	ABK33922	Human immune syste	C 307	34	3.9	339	23	ABV56910	Human prostate exp
C 235	35	4.1	11523	24	ABK28323	DNA transcription	308	34	3.9	340	22	AAI93761	Human polynucleoti
C 236	35	4.1	12601	24	ABL34206	Human immune syste	309	34	3.9	343	22	AAI85587	Human cardiovascul
C 237	35	4.1	12601	24	ABL33999	Human immune syste	310	34	3.9	343	22	AAI86988	Human polynucleoti
C 238	35	4.1	21537	24	AA546788	Tumour suppressor	311	34	3.9	344	22	AAI85033	Human polynucleoti
C 239	35	3.9	61020	22	AA529173	Genomic sequence #	C 312	34	3.9	345	24	ABQ85902	Arabidopsis thalia
240	34	3.9	77	22	AAK7876	Human immune/haema	313	34	3.9	346	22	AAI87279	Human polynucleoti
241	34	3.9	77	22	AAK79578	Human immune/haema	314	34	3.9	356	22	ABV49201	Human prostate exp
242	34	3.9	77	22	AAK84669	Human immune/haema	315	34	3.9	358	14	AAA405102	Human brain expres
243	34	3.9	81	21	AAAC13249	Human secreted pro	316	34	3.9	359	21	AAA40513	Murine adult bone
C 244	34	3.9	82	24	ABN73418	Bovine embryonic g	C 317	34	3.9	360	24	ABQ85688	Arabidopsis thalia
C 245	34	3.9	90	12	AAO11762	Self-complementary	318	34	3.9	360	23	ABV59054	Human prostate exp
C 246	34	3.9	130	22	AA556565	Human CDNA for an	C 319	34	3.9	363	23	ABV12992	Human prostate exp
C 247	34	3.9	139	22	AAI25951	Human breast cance	320	34	3.9	364	23	ABV45789	Human prostate exp
C 248	34	3.9	141	16	AAI06012	Immunodominant fra	321	34	3.9	365	22	AAI83450	Human polynucleoti
C 249	34	3.9	143	23	ABV19092	Human prostate exp	C 322	34	3.9	365	24	ABQ57966	Human colon cancer
250	34	3.9	156	22	AAK63711	Human immune/haema	323	34	3.9	369	18	AAI88086	3' portion of cDNA
251	34	3.9	159	22	AAK56708	Human immune/haema	324	34	3.9	369	18	AAV02169	Human secreted pro
252	34	3.9	160	23	ABV48759	Human prostate exp	325	34	3.9	370	21	AAZ51574	Haematobia irritan
C 253	34	3.9	171	23	ABV57330	Human prostate exp	326	34	3.9	372	21	AAI83172	Human polynucleoti
C 254	34	3.9	174	23	ABV37373	Human prostate exp	327	34	3.9	372	22	AAI83172	Human polynucleoti
C 255	34	3.9	174	24	ABL67899	Human ovarian can	328	34	3.9	373	22	AAI83172	Human polynucleoti
C 256	34	3.9	177	23	ABV55931	Human prostate exp	329	34	3.9	374	22	AA525015	Human ovarian PCR
257	34	3.9	181	23	ABV54276	Human prostate exp	330	34	3.9	374	22	AAH83660	Human ovarian tumo
258	34	3.9	190	22	AAI08206	Human breast cance	331	34	3.9	375	22	AAI90861	Human polynucleoti
259	34	3.9	191	17	AAI13475	Capture probe for	C 332	34	3.9	375	23	ABV15130	Human prostate exp
260	34	3.9	200	23	ABV19076	Human prostate exp	333	34	3.9	375	23	ABV58452	Human prostate exp
C 261	34	3.9	202	22	AA529171	Genomic sequence #	334	34	3.9	376	22	AAI84760	Human polynucleoti
262	34	3.9	202	22	AAK78762	Human immune/haema	335	34	3.9	377	21	AA60071	Human secreted pro
263	34	3.9	202	22	AAK79575	Human immune/haema	336	34	3.9	378	22	AAI90863	Human polynucleoti
264	34	3.9	202	22	AAK84668	Human prostate exp	337	34	3.9	380	22	AAI88350	Human polynucleoti
265	34	3.9	203	23	ABV56674	Human prostate exp	338	34	3.9	382	23	ABV54300	Human prostate exp
266	34	3.9	206	24	ABL6453	Human prostate exp	339	34	3.9	382	23	ABV56122	Human prostate exp
C 267	34	3.9	209	23	ABV07955	Human prostate exp	340	34	3.9	383	23	ABV57015	Human prostate exp
268	34	3.9	212	23	ABV57286	Human prostate exp	341	34	3.9	384	22	AAI85450	Human polynucleoti
269	34	3.9	227	23	ABV59030	Human prostate exp	342	34	3.9	384	22	AAI88641	Human polynucleoti
270	34	3.9	227	23	ABV20080	Human prostate exp	C 343	34	3.9	385	22	AAI13572	Human polynucleoti
271	34	3.9	229	23	ABV57683	Human prostate exp	344	34	3.9	388	22	AAI87805	Human prostate exp
272	34	3.9	230	23	ABV57465	Human prostate exp	345	34	3.9	388	22	AAI89390	Human prostate exp
273	34	3.9	232	23	ABV58523	Human prostate exp	C 346	34	3.9	389	22	ABV10587	Human prostate exp
274	34	3.9	236	23	ABV56953	Human prostate exp	347	34	3.9	389	22	AAI88025	Human polynucleoti
275	34	3.9	239	22	ABAI1191	Human nervous syst	348	34	3.9	389	22	AAI91234	Human polynucleoti
276	34	3.9	240	23	ABV58851	Human prostate exp	349	34	3.9	390	22	AAI86998	Human polynucleoti
277	34	3.9	251	22	AAK88384	Human digestive sy	350	34	3.9	391	22	AAI87310	Human polynucleoti
278	34	3.9	251	22	AAI57596	Human colorectal c	351	34	3.9	393	22	AAI87824	Human polynucleoti
C 279	34	3.9	254	24	ABK64649	Human benign prost	352	34	3.9	393	22	AAI88999	Human polynucleoti
C 280	34	3.9	254	24	ABL66183	Lung cancer relate	353	34	3.9	394	24	AAI87163	Human prostate exp
C 281	34	3.9	254	24	ABL68715	Prostate cancer re	C 354	34	3.9	398	23	ABV34113	Human prostate exp
C 282	34	3.9	268	23	ABV38724	Human prostate exp	C 355	34	3.9	398	23	ABV42975	Human prostate exp
C 283	34	3.9	272	22	AAI12612	Human breast cance	356	34	3.9	400	22	AAI82922	Human polynucleoti
284	34	3.9	274	23	ABV58263	Human prostate exp	357	34	3.9	400	22	AAI87341	Human polynucleoti
285	34	3.9	280	23	ABV58623	Human prostate exp	358	34	3.9	402	22	AAI82881	Human polynucleoti
286	34	3.9	282	23	ABV47775	Human prostate exp	359	34	3.9	403	22	AAI87342	Human polynucleoti
287	34	3.9	287	23	ABV18679	Human prostate exp	360	34	3.9	403	22	AAI893629	Human polynucleoti
288	34	3.9	289	14	AAQ53450	Sequence of the 3'	361	34	3.9	404	22	AAI89005	Human polynucleoti
289	34	3.9	289	21	AAI88427	Human polynucleoti	362	34	3.9	405	22	AAI84799	Human polynucleoti
290	34	3.9	291	21	AAA40414	Potato infection 1	363	34	3.9	405	22	AAI84799	Human polynucleoti
291	34	3.9	291	24	ABK46641	CDNA encoding colo	364	34	3.9	410	23	ABV54332	Human prostate exp
292	34	3.9	291	24	ABK46641	CDNA encoding colo	365	34	3.9	411	22	AAI87998	Human polynucleoti
293	34	3.9	296	22	AAI09995	Human breast cance	366	34	3.9	411	22	AAI88712	Human polynucleoti
294	34	3.9	296	22	AAI19575	Human breast cance	367	34	3.9	412	22	AAI80754	Human polynucleoti
295	34	3.9	304	23	ABV19428	Human prostate exp	368	34	3.9	412	22	AAI87374	Human polynucleoti
296	34	3.9	310	24	ABL87211	Human ovarian can	369	34	3.9	412	23	ABV56553	Human prostate exp
C 297	34	3.9	311	24	ABL87180	Human ovarian can	370	34	3.9	413	22	AAI89034	Human polynucleoti
C 298	34	3.9	312	22	AAK57869	Human immune/haema	371	34	3.9	415	22	AAI85235	Human polynucleoti
299	34	3.9	314	22	ABV58492	Human prostate exp	372	34	3.9	415	22	AAI92380	Human polynucleoti
300	34	3.9	321	21	AAV70159	Plasmodium falcipa	373	34	3.9	416	22	AAI87599	Human polynucleoti
C 301	34	3.9	321	21	AAV70159	Plasmodium falcipa	374	34	3.9	418	22	AAI83742	Human polynucleoti

375	34	3.9	418	22	AA188027	Human polynucleoti	448	34	3.9	556	23	ABV56471	Human prostate exp
376	34	3.9	418	23	ABV49108	Human prostate exp	C 449	34	3.9	556	24	ABO36996	Oligonucleotide fo
377	34	3.9	418	23	ABV58447	Human prostate exp	450	34	3.9	556	24	ABO36997	Oligonucleotide fo
378	34	3.9	420	22	AA189182	Human polynucleoti	C 451	34	3.9	559	24	ABO56948	Human colon cancer
379	34	3.9	421	22	AA189184	Human polynucleoti	452	34	3.9	555	23	ABV59100	Human prostate exp
C 380	34	3.9	422	22	ABAI3982	Human nervous syst	C 453	34	3.9	572	24	AB146689	B tropicalis aller
381	34	3.9	423	22	ABAI82334	Human polynucleoti	C 454	34	3.9	592	24	AB101467	Murine apoptosis r
382	34	3.9	423	22	AA188083	Human polynucleoti	C 455	34	3.9	593	23	ABV04700	Human prostate exp
383	34	3.9	425	22	AAH4746	Human colon cancer	C 456	34	3.9	599	24	ABO66406	Arabidopsis thalia
384	34	3.9	425	22	ABV57920	Human prostate exp	C 457	34	3.9	600	24	ABO58060	Human colon cancer
385	34	3.9	426	22	AA183998	Human polynucleoti	C 458	34	3.9	600	24	ABO52494	Oligonucleotide fo
386	34	3.9	426	22	AA184662	Human prostate exp	459	34	3.9	604	20	AAV09282	Oligonucleotide fo
C 387	34	3.9	427	23	ABV31757	Human prostate exp	C 460	34	3.9	604	21	AAV09282	Cat flea OS- D CDN
C 388	34	3.9	427	23	AB162588	Human prostate exp	461	34	3.9	604	21	AAV09282	Cat flea OS- D CDN
389	34	3.9	428	22	AA185047	Human polynucleoti	C 462	34	3.9	608	21	AAV09282	Human secreted pro
390	34	3.9	431	22	AA183653	Human polynucleoti	C 463	34	3.9	608	21	AAV09282	Human secreted pro
391	34	3.9	431	22	AA189345	Human polynucleoti	464	34	3.9	612	22	AAH71471	Arabidopsis thalia
392	34	3.9	433	22	ABV57898	Human prostate exp	465	34	3.9	612	22	AAH71471	Human cervical can
393	34	3.9	435	22	AA188171	Human polynucleoti	466	34	3.9	615	20	AAV19927	DNA encoding a hum
394	34	3.9	437	22	ABV01418	Human prostate exp	C 467	34	3.9	631	19	AAV09282	SSG #5. Homo sapi
C 395	34	3.9	437	22	ABV01418	Human prostate exp	C 468	34	3.9	633	19	AAV09282	Nucleotide sequenc
396	34	3.9	441	22	AA186818	Human polynucleoti	469	34	3.9	633	19	AAV09282	Human secreted pro
397	34	3.9	443	22	AA183563	Human polynucleoti	470	34	3.9	638	20	AAV09282	DNA encoding a hum
C 398	34	3.9	444	22	AA191819	Human prostate exp	471	34	3.9	640	21	AAV09282	Human secreted pro
C 399	34	3.9	445	22	ABV37681	Human polynucleoti	C 472	34	3.9	640	21	AAV09282	Oligonucleotide fo
C 400	34	3.9	446	23	ABV48854	Human prostate exp	473	34	3.9	642	24	ABO53738	Oligonucleotide fo
C 401	34	3.9	446	24	AB194019	Arabidopsis thalia	C 474	34	3.9	646	21	AAV09282	CDNA sequence of a
C 402	34	3.9	447	22	AA181512	Human polynucleoti	C 475	34	3.9	646	24	ABO66002	Human prostate exp
C 403	34	3.9	451	23	ABV14959	Human prostate exp	C 476	34	3.9	647	23	ABV48317	Human colon cancer
C 404	34	3.9	451	22	AA183042	Human polynucleoti	C 477	34	3.9	650	24	ABO60388	Human prostate exp
405	34	3.9	458	21	AAZ07549	Maize glutathione-	C 478	34	3.9	655	22	AAH70113	Human cervical can
406	34	3.9	458	21	AAZ07549	Maize glutathione-	C 479	34	3.9	660	20	AAH37406	Human secreted pro
407	34	3.9	458	21	AAZ94789	Maize class III gl	C 480	34	3.9	665	22	AAH34083	Human colon cancer
C 408	34	3.9	458	21	AAZ94789	Maize class III gl	C 481	34	3.9	671	22	AAV05318	Human secreted pro
C 409	34	3.9	462	22	AA184644	Human polynucleoti	C 482	34	3.9	678	22	AAV05318	Human secreted pro
410	34	3.9	465	23	ABV55314	Human prostate exp	C 483	34	3.9	693	24	ABO32104	Oligonucleotide fo
411	34	3.9	465	23	ABV55314	Human prostate exp	C 484	34	3.9	703	24	AA142577	Partial sequence o
C 412	34	3.9	466	22	AA188546	Human polynucleoti	C 485	34	3.9	709	24	AAV08497	Human extracellular
C 413	34	3.9	468	23	ABV07439	Human prostate exp	486	34	3.9	728	21	AAV08497	Human ORFX ORF2764
C 414	34	3.9	470	23	ABV45926	Human prostate exp	487	34	3.9	732	22	AAH34494	Human colon cancer
C 415	34	3.9	470	23	ABV45926	Human prostate exp	C 488	34	3.9	738	22	AAH34494	Human cDNA SEQ ID
416	34	3.9	470	23	ABV48080	Human prostate exp	C 489	34	3.9	738	22	AAH34494	Human breast cancer
C 417	34	3.9	474	11	AAO03413	Nucleotide sequenc	C 490	34	3.9	738	22	AAH34494	Human immunoglobul
C 418	34	3.9	474	11	AAO03413	Nucleotide sequenc	C 491	34	3.9	744	21	AAH34494	Human secreted pro
C 419	34	3.9	474	23	ABV56492	Human prostate exp	C 492	34	3.9	744	21	AAH34494	Human secreted pro
C 420	34	3.9	475	22	AA108177	Human breast cance	493	34	3.9	746	20	AAH34494	CDNA encoding a mu
421	34	3.9	478	22	AA193637	Human polynucleoti	494	34	3.9	748	12	AAO10376	DNA encoding human
422	34	3.9	478	22	AA193637	Human polynucleoti	495	34	3.9	748	12	AAO10376	Human neuroblastom
C 423	34	3.9	479	24	ABO60200	Human colon cancer	C 496	34	3.9	756	22	AAH33304	Human secreted pro
424	34	3.9	481	22	AA187727	Human polynucleoti	497	34	3.9	760	20	AAV49425	Human secreted pro
C 425	34	3.9	482	23	ABV56622	Human prostate exp	498	34	3.9	760	20	AAV49425	Human secreted pro
C 426	34	3.9	482	23	ABV56622	Human prostate exp	C 499	34	3.9	761	24	ABO32389	Oligonucleotide fo
427	34	3.9	486	22	AAK64568	Human immune/haema	500	34	3.9	761	24	ABO32389	Oligonucleotide fo
C 428	34	3.9	487	21	AAK74174	Human immune/haema	501	34	3.9	764	21	AAV1707	Human breast and o
C 429	34	3.9	501	21	AAAF09443	Fusarium venenatum	502	34	3.9	765	21	AAV1707	Human secreted pro
430	34	3.9	502	23	ABV56360	Human prostate exp	503	34	3.9	776	24	AB180464	Human polynucleoti
431	34	3.9	504	23	ABV57520	Human prostate exp	C 504	34	3.9	776	24	AB180464	Human polynucleoti
432	34	3.9	505	23	ABV48870	Human prostate exp	505	34	3.9	785	22	AA195872	Human neuroblastom
433	34	3.9	507	23	ABV53326	Human prostate exp	C 506	34	3.9	798	13	AAQ27948	Human neuroblastom
434	34	3.9	511	20	AAZ4402	Human prostate exp	507	34	3.9	805	22	AAH64808	Human secreted pro
C 435	34	3.9	520	23	ABV54298	Human prostate exp	508	34	3.9	815	21	AAV09282	Human secreted pro
C 436	34	3.9	521	23	ABV60727	Human prostate exp	509	34	3.9	816	21	AAV09282	Human secreted pro
437	34	3.9	521	23	ABV47378	Human prostate exp	C 510	34	3.9	837	24	ABO3188	Oligonucleotide fo
438	34	3.9	522	23	AAH34775	Human colon cancer	511	34	3.9	837	24	ABO3188	Oligonucleotide fo
C 439	34	3.9	536	23	ABV37538	Human prostate exp	C 512	34	3.9	843	23	ABV16074	Human prostate exp
C 440	34	3.9	536	23	ABO35232	Oligonucleotide fo	513	34	3.9	846	19	AAV49819	Human polynucleoti
C 441	34	3.9	536	24	ABO35232	Oligonucleotide fo	514	34	3.9	861	19	AAV49819	Human polynucleoti
442	34	3.9	546	24	ABO35232	CDNA encoding nove	C 515	34	3.9	868	24	ABO56024	Human ovarian anti
443	34	3.9	546	24	ABO35232	CDNA encoding nove	516	34	3.9	877	24	ABO56024	Oligonucleotide fo
444	34	3.9	556	21	AAZ52551	Human secreted pro	C 517	34	3.9	877	24	ABO36215	Oligonucleotide fo
445	34	3.9	556	23	ABV40063	Human prostate exp	518	34	3.9	878	22	AAK64702	Human immune/haema
446	34	3.9	556	23	ABV40163	Human prostate exp	519	34	3.9	898	22	AAH33229	DNA encoding human
447	34	3.9	556	23	ABV42105	Human prostate exp	520	34	3.9	899	24	AB190124	Human polynucleoti
448	34	3.9	556	23	ABV43601	Human prostate exp							





667	34	3	9	1694	22	AAS25837	Human cDNA encoding
668	34	3	9	1696	24	ABO55028	Human ovarian anti
669	34	3	9	1697	21	AAC77781	Human cancer assoc
670	34	3	9	1709	21	AAZ35180	Corn delta-9 stear
671	34	3	9	1710	14	AAQ47966	Rape acyl-ACP thio
672	34	3	9	1710	21	AAF21694	Human breast and o
673	34	3	9	1710	23	ABV22936	Human prostate exp
674	34	3	9	1710	23	ABV25034	Human prostate exp
675	34	3	9	1710	23	ABV25216	Human prostate exp
676	34	3	9	1710	23	ABV2591	Human prostate exp
677	34	3	9	1710	23	ABV25843	Human prostate exp
678	34	3	9	1710	23	ABV28767	Human prostate exp
679	34	3	9	1729	22	ABA06527	Human prostate exp
680	34	3	9	1747	20	AAK97979	Human CDNA, SEQ ID
681	34	3	9	1749	22	AAI60696	Human secreted pro
682	34	3	9	1749	24	ABL90045	Human polynucleoti
683	34	3	9	1754	22	AAS33356	DNA encoding huma
684	34	3	9	1760	21	AAC59930	Human secreted pro
685	34	3	9	1760	21	AAZ43782	Human fetal brain
686	34	3	9	1784	21	AAZ33337	Human secreted pro
687	34	3	9	1792	21	AAF15784	Human prostate can
688	34	3	9	1816	22	AAF25489	Nucleotide sequenc
689	34	3	9	1823	22	AAS29041	CNA encoding for
690	34	3	9	1827	24	ABK88010	CNA encoding huma
691	34	3	9	1831	21	AAA64634	CNA encoding huma
692	34	3	9	1836	17	AAV33461	Human T1-receptor
693	34	3	9	1837	17	AAAT33852	Rat vas deferens p
694	34	3	9	1841	16	AAQ91311	Human Fas-L gene,
695	34	3	9	1841	20	AAK24879	Human Fas ligand (
696	34	3	9	1841	20	AAK21574	Human FasL protein
697	34	3	9	1852	21	AACT77519	Human ORFX, ORF3074
698	34	3	9	1858	18	AAAT72175	Alzheimer's disease
699	34	3	9	1866	20	AAV84466	Human secreted pro
700	34	3	9	1866	22	ABA83249	Human secreted pro
701	34	3	9	1875	20	AAZ20076	Human leupaxin CDN
702	34	3	9	1902	24	ABA96106	Human mut dGTPase
703	34	3	9	1913	21	AAC59920	Human secreted pro
704	34	3	9	1918	22	AAAD05399	Human secreted pro
705	34	3	9	1934	17	AAAT13229	OR-1 orphan recept
706	34	3	9	1946	21	AAC988332	Human colon cancer
707	34	3	9	1963	24	ABL89766	Human polynucleoti
708	34	3	9	1964	22	AAH15328	Human cDNA, sequenc
709	34	3	9	1971	20	AAK08428	Acidic leucine ami
710	34	3	9	1976	23	ABV22114	Human prostate exp
711	34	3	9	1978	14	AAQ37189	Petunia, b5P synth
712	34	3	9	1999	23	ABL28468	Drosophila, melanog
713	34	3	9	2021	24	ABO54639	Human ovarian anti
714	34	3	9	2030	20	AAZ23135	Wheat sucrose tran
715	34	3	9	2034	22	AAAD06011	Human neuronal apo
716	34	3	9	2060	24	AAD33348	Human consensus se
717	34	3	9	2092	24	ABL90358	Human polynucleoti
718	34	3	9	2094	21	AAA39067	Human secreted pro
719	34	3	9	2106	22	AAH33397	Human colon cancer
720	34	3	9	2123	24	ABO35071	Human ovarian anti
721	34	3	9	2143	16	AAQ90242	Full length Sorghu
722	34	3	9	2159	22	AAF98708	Human ovarian canc
723	34	3	9	2164	21	AAK75653	Human ORFX, ORF1208
724	34	3	9	2187	21	AAZ92220	Human thioedoxin
725	34	3	9	2187	22	AAAD05580	Human secreted pro
726	34	3	9	2187	24	AAIT72961	Human thioedoxin
727</							

813	34	3.9	4419	22	AA006838	Human STEAP-3 cDNA	C 886	34	3.9	6255	24	ABL32960	Human immune syste
814	34	3.9	4429	22	AA007073	Human six transmem	C 887	34	3.9	6338	24	ABL33126	Human immune syste
815	34	3.9	4451	22	ABN89471	Human G-protein co	C 888	34	3.9	6350	24	ABL70499	Chemically treated
C 816	34	3.9	4626	23	ABL27960	Drosophila melanog	C 889	34	3.9	6350	24	ABL34628	Human metastasis a
C 817	34	3.9	4654	22	AA546781	Tumour suppressor	C 890	34	3.9	6407	24	ABL32942	Human immune syste
C 818	34	3.9	4654	24	ABL34223	Human immune syste	C 891	34	3.9	6430	24	ABN80122	Human chemically m
C 819	34	3.9	5173	18	ABL89783	Human immunodefici	C 892	34	3.9	6436	24	ABL33693	Human immune syste
C 820	34	3.9	5173	24	ABN96859	Gene #3357 used to	C 893	34	3.9	6436	24	ABL33409	Signal transductio
C 821	34	3.9	5193	21	AACT6457	Human ORFX ORF2012	C 894	34	3.9	6476	24	ABL32540	Human immune syste
C 822	34	3.9	5198	24	ABL32279	Human immune syste	C 895	34	3.9	6485	22	AA546559	Tumour suppressor
C 823	34	3.9	5267	20	AA021357	Human BAI3 gene.	C 896	34	3.9	6485	24	ABL33088	Human immune syste
C 824	34	3.9	5304	24	ABK28356	DNA transcription	C 897	34	3.9	6522	24	ABL33050	Human immune syste
C 825	34	3.9	5310	24	ABL70418	Chemically treated	C 898	34	3.9	6534	22	AA545404	Chemically pretrea
C 826	34	3.9	5310	24	ABK31445	Signal transductio	C 899	34	3.9	6534	24	ABN80184	Human chemically m
C 827	34	3.9	5314	24	ABL54310	Chemically treated	C 900	34	3.9	6534	24	ABL33426	Human immune syste
C 828	34	3.9	5314	24	ABL32161	Human immune syste	C 901	34	3.9	6534	24	ABK28251	DNA transcription
C 829	34	3.9	5368	24	ABN80099	Human chemically m	C 902	34	3.9	6668	24	ABL33696	Human immune syste
C 830	34	3.9	5386	24	ABL32533	Human immune syste	C 903	34	3.9	6794	24	ABL70219	Chemically treated
C 831	34	3.9	5437	24	ABL33742	Human immune syste	C 904	34	3.9	6794	24	AA561174	Human gene regulat
C 832	34	3.9	5539	24	ABL34626	Human metastasis a	C 905	34	3.9	6794	24	ABK31264	Signal transductio
C 833	34	3.9	5586	24	ABK40004	Human chemically p	C 906	34	3.9	6802	24	ABL33231	Human immune syste
C 834	34	3.9	5586	24	ABL32649	Human immune syste	C 907	34	3.9	6802	24	ABL33894	Human immune syste
C 835	34	3.9	5739	24	ABL32719	Human immune syste	C 908	34	3.9	6922	22	AA546534	Tumour suppressor
C 836	34	3.9	5752	22	AA85087	Nucleotide sequenc	C 909	34	3.9	6922	24	ABL70380	Chemically treated
C 837	34	3.9	5771	22	ABN80073	Human chemically m	C 910	34	3.9	6922	24	AA561332	Human gene regulat
C 838	34	3.9	5828	20	AA084552	AML1-MTG16 fusion	C 911	34	3.9	6922	24	ABK31411	Signal transductio
C 839	34	3.9	5883	24	ABL32353	Human immune syste	C 912	34	3.9	6923	24	ABL70275	Chemically treated
C 840	34	3.9	5883	24	ABL34473	Human metastasis a	C 913	34	3.9	6923	24	AA561198	Human gene regulat
C 841	34	3.9	5937	24	ABN80150	Human chemically m	C 914	34	3.9	6923	24	ABK31296	Signal transductio
C 842	34	3.9	5937	24	ABL34542	Human metastasis a	C 915	34	3.9	6960	24	ABL34399	Human immune syste
C 843	34	3.9	5956	24	ABL70325	Chemically treated	C 916	34	3.9	7061	24	ABL70247	Chemically treated
C 844	34	3.9	5956	24	AA561272	Human gene regulat	C 917	34	3.9	7061	24	ABL32996	Human immune syste
C 845	34	3.9	5956	24	ABK31368	Signal transductio	C 918	34	3.9	7061	24	AA561188	Human gene regulat
C 846	34	3.9	5963	24	ABL34612	Human metastasis a	C 919	34	3.9	7070	24	ABN80054	Human chemically m
C 847	34	3.9	5975	24	ABN80064	Human chemically m	C 920	34	3.9	7168	24	ABL92231	Chemically treated
C 848	34	3.9	5990	24	ABK52570	Rat sequence diffe	C 921	34	3.9	7168	24	ABD22318	Chemically treated
C 849	34	3.9	6002	20	AAK60227	Poly(A) binding pr	C 922	34	3.9	7302	24	ABL32710	Human immune syste
C 850	34	3.9	6012	24	ABL34058	Human immune syste	C 923	34	3.9	7302	24	ABL32345	Human immune syste
C 851	34	3.9	6024	24	ABL33647	Human immune syste	C 924	34	3.9	7332	24	ABK33924	Human DNA for stag
C 852	34	3.9	6027	24	ABN80172	Human chemically m	C 925	34	3.9	7432	24	ABL32167	Human immune syste
C 853	34	3.9	6035	24	ABL33524	Human immune syste	C 926	34	3.9	7432	24	AAAD28386	Human chemically t
C 854	34	3.9	6036	24	ABN80117	Human chemically m	C 927	34	3.9	7455	24	ABL33759	Human immune syste
C 855	34	3.9	6042	24	ABL33593	Human immune syste	C 928	34	3.9	7588	22	AA545484	Chemically pretrea
C 856	34	3.9	6044	24	ABL32840	Human immune syste	C 929	34	3.9	7588	24	ABL34013	Human immune syste
C 857	34	3.9	6056	20	AA084593	AML1-MTG16 fusion	C 930	34	3.9	7588	24	ABK28410	DNA transcription
C 858	34	3.9	6057	24	ABL70361	Chemically treated	C 931	34	3.9	7603	24	ABL54329	Chemically treated
C 859	34	3.9	6057	24	ABK31396	Signal transductio	C 932	34	3.9	7603	24	ABL32308	Human immune syste
C 860	34	3.9	6059	24	ABL32812	Human immune syste	C 933	34	3.9	7657	22	AA545477	Chemically pretrea
C 861	34	3.9	6063	22	AA546338	Tumour suppressor	C 934	34	3.9	7657	24	ABL34022	Human immune syste
C 862	34	3.9	6070	24	ABO67130	Human angiogenesis	C 935	34	3.9	7749	24	ABL70436	Chemically treated
C 863	34	3.9	6070	24	ABL70372	Chemically treated	C 936	34	3.9	7781	24	ABL33168	Human immune syste
C 864	34	3.9	6070	24	ABL92199	Chemically treated	C 937	34	3.9	7900	22	AAK69080	Human immune/haema
C 865	34	3.9	6070	24	ABL49310	Human polynucleoti	C 938	34	3.9	7906	22	AA545391	Chemically pretrea
C 866	34	3.9	6070	24	ABL32241	Human immune syste	C 939	34	3.9	7906	22	ABN80179	Human chemically m
C 867	34	3.9	6070	24	ABL34579	Human immune syste	C 940	34	3.9	7906	24	ABK28236	DNA transcription
C 868	34	3.9	6070	24	ABL34579	Human metastasis a	C 941	34	3.9	7913	22	AA546312	Tumour suppressor
C 869	34	3.9	6090	22	AA546316	Tumour suppressor	C 942	34	3.9	7913	24	ABL70532	Chemically treated
C 870	34	3.9	6090	24	ABL70184	Chemically treated	C 943	34	3.9	7913	24	AA561129	Human gene regulat
C 871	34	3.9	6090	24	ABL32577	Human immune syste	C 944	34	3.9	7913	24	ABK31217	Signal transductio
C 872	34	3.9	6090	24	ABL33578	Human immune syste	C 945	34	3.9	7925	24	ABN80130	Human chemically m
C 873	34	3.9	6107	24	ABN80104	Human chemically m	C 946	34	3.9	8186	24	ABN80090	Human chemically m
C 874	34	3.9	6115	24	ABL33801	Human immune syste	C 947	34	3.9	8234	24	ABL33649	Human immune syste
C 875	34	3.9	6137	24	ABL70128	Chemically treated	C 948	34	3.9	8239	18	AAAT61981	Human ataxia and r
C 876	34	3.9	6137	24	ABL34451	Human metastasis a	C 949	34	3.9	8254	24	ABL32864	Human immune syste
C 877	34	3.9	6145	24	ABL32973	Human immune syste	C 950	34	3.9	8298	24	ABN80191	Human chemically m
C 878	34	3.9	6162	24	ABK31331	Signal transductio	C 951	34	3.9	8305	24	ABL33568	Human immune syste
C 879	34	3.9	6171	24	ABL32789	Human immune syste	C 952	34	3.9	8342	24	ABL32501	Human immune syste
C 880	34	3.9	6172	24	ABN80042	Human chemically m	C 953	34	3.9	8404	22	AA546500	Tumour suppressor
C 881	34	3.9	6172	24	ABL34573	Human metastasis a	C 954	34	3.9	8404	24	ABL33595	Human immune syste
C 882	34	3.9	6196	24	ABL33864	Human immune syste	C 955	34	3.9	8607	24	ABL92303	Chemically treated
C 883	34	3.9	6211	24	ABL32806	Human immune syste	C 956	34	3.9	8607	24	ABL49372	Human polynucleoti
C 884	34	3.9	6240	24	ABL32136	Human immune syste	C 957	34	3.9	8771	24	ABL33825	Human immune syste
C 885	34	3.9	6240	24	ABL34454	Human metastasis a	C 958	34	3.9	8952	22	AA546446	Tumour suppressor

	Human DNA for strand
	Human chemically P
	Human immune system
	Human angiotensinase II
	Human immune system
	Human immune system
	Human immune system
	Chemically pretreated
	Human chemically m
	Human immune system
	Human immune system
	Genomic sequence #
	Chemically treated
	Human immune system
	Chemically pretreated
	DNA transcription
	DNA transcription
	Human immune system
	Chemically treated
	DNA transcription
	Human immune system
	Human immune system
	Human immune system
	DNA transcription
	Human immune system
	Human immune system
	Human immune system
	Chemically pretreated
	Tumour suppressor
	Human immune system
	Human immune system
	Human chemically P
	Chemically pretreated
	Human OSR31 genome
	Tumour suppressor
	Chemically pretreated
	Tumour suppressor
	Human immune system
	DNA transcription
	Human immune system

- inflammatory;
- metabolic; AIDS;
- arteriosclerosis;
- hypothyroidism;
- inactivity;
- arthritis;

29-NOV-2001.  
25-MAY-2001; 2001WO-US17150.  
25-MAY-2000; 2000US-207901P.  
01-JUN-2000; 2000US-208983P.  
07-JUN-2000; 2000US-209861P.  
15-JUN-2000; 2000US-211825P.  
22-JUN-2000; 2000US-213744P.

(INCY-) INCYTE GENOMICS INC.

Yue H, Sanjarnwala MS, Baugm MR, Gandhi AR, Ring HZ, Elliott V, Wallia NK, Yang M, Khan FA, Ramkumar J, Tang YT, Hafalia A, Lal P, Nguyen DB, Yang MG, Lee EA, Tribouley CM, Patterson C, Lu Y, Buford N, Ding L, Bruns CM, Kearney L, Reddy R.

WPI; 2002-097650/13.

New human drug metabolizing enzymes and polynucleotides encoding the enzyme for diagnosing, preventing or treating cell proliferative, autoimmune/inflammatory, endocrine, eye, metabolic and gastrointestinal disorders -

Claim 5; Page 154-155; 158pp; English.

The present sequence is human drug metabolizing enzyme (DME)-8 cDNA. DME polypeptide, polynucleotide and modulators are useful for diagnosis, treatment and prevention of autoimmune/inflammatory, cell proliferative, developmental, endocrine, eye, metabolic, and gastrointestinal disorders, including liver disorders.

The autoimmune/inflammatory disorders treatable include AIDS, adult respiratory distress syndrome, Addison's disease, allergies, anaemia, asthma, atherosclerosis, osteoporosis, autoimmune haemolytic anaemia, autoimmune thyroiditis, Crohn's disease, atrophic dermatitis, diabetic mellitus, Graves' disease, glomerulonephritis, rheumatoid arthritis, scleroderma, systemic lupus erythematosus, systemic sclerosis, ulcerative colitis, haemodialysis and ureitis, viral, bacterial, fungal, parasitic, protozoal, helminthic infections and trauma, and cell proliferative disorders such as cancer, actinic keratosis, arteriosclerosis, atherosclerosis, burns, cirrhosis, hepatitis and psoriasis. Developmental disorders include anaemia, renal tubular acidosis, epilepsy, hypothyroidism and cataract, and endocrine disorders include disorders of hypothalamus and pituitary, disorders associated with hypopituitarism, including sarcoidosis, diabetes insipidus, hypogonadism, disorders associated with hypothyroidism including goitre, acute thyroiditis, Graves' disease, disorders associated with hyperparathyroidism, pancreatic disorders such as type I or type II diabetes mellitus, disorders associated with adrenals such as hyperplasia, Cushing's disease, endometriosis, infertility, hypergonadal disorders, and gynaecomastia. Eye disorders include conjunctivitis, keratitis, glaucoma and macular degeneration, and metabolic disorders include diabetes, cystic fibrosis, goitre, hypercholesterolaemia, hypoglycaemia, hyperlipidaemia, lysosomal storage diseases, obesity, phenylketonuria and hypocalcaemia. Also the molecules are useful for treating gastrointestinal disorders such as dysphagia, gastritis, peptic ulcer, cholelithiasis, cirrhosis, hepatitis, hyperbilirubinaemia, constipation, diarrhoea, jaundice, Wilson's disease, thrombosis and hepatic tumours. The DME polypeptide is also useful for screening its agonist or antagonist.

Sequence 2037 BP; 427 A; 589 C; 582 G; 439 T; 0 other;

4.8%; Score 41; DB 24; Length 2037;

```

      41; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

```

823 TTTATATATAAAAAAAAAAAAAAAAAAAAAAAAAA 863

1935 TTTATATATAAAAAAAAAAAAAAAAAAAAAAAAAA 1975

```

RESULT 2
AAV44318
ID AAV44318 standard; cDNA to mRNA; 2307 BP.
XX
AC AAV44318;
XX
DT 07-OCT-1998 (first entry)
XX
DE Maize starch-associated protein cDNA.
XX
KM Starch-associated protein; maize; starch grain; soluble;
XX
KW transgenic plant; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 33..1946
FT /tag= a
FT /product= "starch-associated protein"
FT /note= "no start codon given"
XX
PN DE19653176-A1.
XX
PD 25-JUN-1998.
XX
PF 19-DEC-1996; 96DE-1053176.
XX
PR 19-DEC-1996; 96DE-1053176.
XX
PA (PLAN-) PLANTTEC BIOTECHNOLOGIE GMBH.
XX
PI Emmermann M, Kosmann J;
XX
DR WPI; 1998-349465/31.
XX
P-PSDB; AAW64231.
XX
PT DNA encoding maize starch-associated protein in potatoes - and
PT corresponding protein, vectors, transgenic plants, antibody, starch,
PT etc.
XX
PS Claim 1; Page 43-46; 55pp; German.
XX
CC This sequence encodes a novel maize starch-associated protein that
CC occurs both bound to starch grains and in soluble form in plant cells.
CC Such a protein can be used to produce transgenic plants or plant cells
CC which are capable of synthesising modified starch.
XX
CC
XX
SQ Sequence 2307 BP; 693 A; 436 C; 555 G; 623 T; 0 other;
XX
Query Match 4.8%; Score 41; DB 19; Length 2307;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 823 TTTATATATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 863
DB 2201 TTTATATATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2241
XX
RESULT 3
AAV44319
ID AAV44319 standard; DNA; 4329 BP.
XX
AC AAV44319;
XX
DT 07-OCT-1998 (first entry)
XX
DE Seq ID #7 from DE196533176.
XX
KM Starch-associated protein; maize; starch grain; soluble;
XX
KW transgenic plant; ss.
XX
OS Unknown.

```

```

XX
PN DE19653176-A1.
XX
PD 25-JUN-1998.
XX
PF 19-DEC-1996; 96DE-1053176.
XX
PR 19-DEC-1996; 96DE-1053176.
XX
PA (PLAN-) PLANTTEC BIOTECHNOLOGIE GMBH.
XX
PI Emmermann M, Kosmann J;
XX
DR WPI; 1998-349465/31.
XX
PT DNA encoding maize starch-associated protein in potatoes - and
PT corresponding protein, vectors, transgenic plants, antibody, starch,
PT etc.
XX
PS Disclosure; Page -; 55pp; German.
XX
CC This specification describes potato and maize starch-associated proteins
CC that occur both bound to starch grains and in soluble form in plant
CC cells. Such a protein can be used to produce transgenic plants or plant
CC cells which are capable of synthesising modified starch.
CC NOTE: This sequence is not described or represented in the specification
CC but is presented in the sequence listing.
XX
SQ Sequence 4329 BP; 1349 A; 791 C; 1048 G; 1141 T; 0 other;
XX
Query Match 4.8%; Score 41; DB 19; Length 4329;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 823 TTTATATATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 863
DB 4267 TTTATATATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4307
XX
RESULT 4
ABN80207/c
ID ABN80207 standard; DNA; 10825 BP.
XX
AC ABN80207;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human chemically modified disease associated gene SEQ ID NO 224.
XX
KM Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
KM heart disease; epilepsy; histone deacetylation; muscular dystrophy;
KM dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
KM antidiabetic; cytostatic; anticonvulsant; ds.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
PN WO200200927-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-BP07536.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIC-) EPICENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130908/17.
XX
PT Novel nucleic acid useful for diagnosis and therapy of diseases

```

















GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2003, 09:39:00 ; Search time 53 Seconds

(without alignments)  
4993.624 Million cell updates/sec

Title: US-10-021-811-35

Perfect score: 863  
Sequence: 1 gccacgagctctatcacacac.....aaaaaaaaaaaaaaaaaaaaa 863

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 15338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 862724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

1: Issued Patents\_NA.\*  
2: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	4.3	1137	1	US-08-706-214-2
2	37	4.3	1856	1	US-08-706-214-1
3	37	4.3	5108	1	US-07-642-002-1
4	37	4.3	10136	1	US-08-353-700-2
5	37	4.3	10136	5	PCR-US95-16216-2
6	36	4.2	1007	4	US-09-171-156A-61
7	36	4.2	1007	4	US-09-171-156A-63
8	36	4.2	1600	4	US-07-861-458C-37
9	36	4.2	1607	6	5196333-3
10	36	4.2	1736	3	US-09-182-81E-22
11	36	4.2	1736	3	US-09-182-81E-24
12	36	4.2	1736	3	US-09-471-528-22
13	36	4.2	1736	3	US-09-471-528-24
14	36	4.2	1736	4	US-09-634-530-22
15	36	4.2	1736	4	US-09-634-530-24
16	36	4.2	19557	5	PCR-US92-06300-1
17	35	4.1	55	1	US-08-113-646A-41
18	35	4.1	570	3	US-08-688-988-44
19	35	4.1	570	1	US-07-885-970A-10
20	35	4.1	570	1	US-08-298-687A-10
21	35	4.1	570	1	US-08-298-829-10
22	35	4.1	609	1	US-08-530-797-9
23	35	4.1	609	2	US-08-787-335-9
24	35	4.1	760	1	US-08-276-452A-49
25	35	4.1	760	2	US-08-798-744-49
26	35	4.1	1020	2	US-09-425-578-5
27	35	4.1	1160	3	US-08-995-159-1

28	35	4.1	1740	1	US-09-362-512A-3	Sequence 3, Appli
29	35	4.1	1740	4	US-08-964-939-3	Sequence 3, Appli
30	35	4.1	1878	4	US-09-465-558-39	Sequence 39, Appli
31	35	4.1	1931	3	US-09-019-942-2	Sequence 2, Appli
32	35	4.1	1931	4	US-09-099-041A-1	Sequence 1, Appli
33	35	4.1	1931	4	US-09-245-281-1	Sequence 1, Appli
34	35	4.1	1931	4	US-09-470-271-2	Sequence 2, Appli
35	35	4.1	1931	4	US-09-207-359B-1	Sequence 1, Appli
36	35	4.1	2007	3	US-08-747-221B-36	Sequence 36, Appli
37	35	4.1	2007	3	US-08-747-221B-38	Sequence 38, Appli
38	35	4.1	2007	4	US-09-005-051-36	Sequence 36, Appli
39	35	4.1	2007	4	US-09-005-051-38	Sequence 38, Appli
40	35	4.1	2908	4	US-09-930-181-1	Sequence 1, Appli
41	34	3.9	141	3	US-08-737-078A-1	Sequence 1, Appli
42	34	3.9	141	5	PCR-US94-04706-1	Sequence 1, Appli
43	34	3.9	289	1	US-08-341-568-3	Sequence 3, Appli
44	34	3.9	289	2	US-08-911-020-3	Sequence 3, Appli
45	34	3.9	370	4	US-09-376-113-1	Sequence 1, Appli
46	34	3.9	458	2	US-08-924-759-7	Sequence 7, Appli
47	34	3.9	458	3	US-09-248-335-7	Sequence 7, Appli
48	34	3.9	487	4	US-09-257-179-22	Sequence 22, Appli
49	34	3.9	633	1	US-08-664-596B-21	Sequence 21, Appli
50	34	3.9	746	4	US-09-013-810-1	Sequence 1, Appli
51	34	3.9	748	1	US-08-361-467B-3	Sequence 3, Appli
52	34	3.9	748	1	US-08-484-332C-3	Sequence 3, Appli
53	34	3.9	796	1	US-08-104-073-2	Sequence 2, Appli
54	34	3.9	900	3	US-09-248-335-61	Sequence 61, Appli
55	34	3.9	903	5	PCR-US95-06406A-21	Sequence 21, Appli
56	34	3.9	1112	4	US-09-173-300-48	Sequence 48, Appli
57	34	3.9	1146	4	US-08-893-654B-5	Sequence 5, Appli
58	34	3.9	1154	3	US-08-651-136C-7	Sequence 7, Appli
59	34	3.9	1154	4	US-09-229-911A-7	Sequence 7, Appli
60	34	3.9	1213	2	US-08-794-796-1	Sequence 1, Appli
61	34	3.9	1213	1	US-08-276-151-6	Sequence 6, Appli
62	34	3.9	1370	4	US-09-026-408-12	Sequence 12, Appli
63	34	3.9	1371	4	US-09-026-408-11	Sequence 1, Appli
64	34	3.9	1490	4	US-08-964-127-3	Sequence 3, Appli
65	34	3.9	1490	4	US-09-496-692-3	Sequence 3, Appli
66	34	3.9	1507	4	US-09-453-323-1	Sequence 1, Appli
67	34	3.9	1558	1	US-08-455-550-7	Sequence 7, Appli
68	34	3.9	1577	4	US-08-821-994-59	Sequence 59, Appli
69	34	3.9	1613	4	US-08-965-048-3	Sequence 3, Appli
70	34	3.9	1641	4	US-08-300-903A-8	Sequence 8, Appli
71	34	3.9	1837	2	US-08-750-134A-4	Sequence 4, Appli
72	34	3.9	1837	4	US-09-363-745-4	Sequence 4, Appli
73	34	3.9	1841	5	PCR-US95-00362-1	Sequence 11, Appli
74	34	3.9	1858	4	US-08-909-965C-11	Sequence 1, Appli
75	34	3.9	1934	4	US-08-776-844-1	Sequence 4, Appli
76	34	3.9	1971	2	US-08-892-770-4	Sequence 2, Appli
77	34	3.9	1971	2	5188642-3	Sequence 1, Appli
78	34	3.9	2143	2	US-08-656-177A-1	Sequence 1, Appli
79	34	3.9	2143	3	US-09-256-797-1	Sequence 1, Appli
80	34	3.9	2187	4	US-09-127-219B-2	Sequence 2, Appli
81	34	3.9	2760	1	US-08-101-593-3	Sequence 1, Appli
82	34	3.9	2760	1	US-08-101-593-3	Sequence 3, Appli
83	34	3.9	2823	1	US-08-398-008A-1	Sequence 1, Appli
84	34	3.9	2823	2	US-08-893-333-1	Sequence 1, Appli
85	34	3.9	2880	4	US-09-235-451-1	Sequence 1, Appli
86	34	3.9	3667	4	US-09-511-625B-1	Sequence 1, Appli
87	34	3.9	4134	4	US-09-162-021B-1	Sequence 1, Appli
88	34	3.9	4134	4	US-09-687-477-17	Sequence 17, Appli
89	34	3.9	5173	4	US-08-242-677-1	Sequence 1, Appli
90	34	3.8	60	4	US-09-884-627-14	Sequence 14, Appli
91	34	3.8	80	4	US-09-884-627-15	Sequence 15, Appli
92	34	3.8	105	4	US-09-284-627-23	Sequence 23, Appli
93	34	3.8	114	1	US-08-120-827-99	Sequence 99, Appli
94	34	3.8	114	1	US-08-478-675-99	Sequence 99, Appli
95	33	3.8	120	1	US-08-153-051B-28	Sequence 28, Appli
96	33	3.8	120	1	US-08-060-952C-44	Sequence 44, Appli
97	33	3.8	120	2	US-08-151-477A-28	Sequence 28, Appli
98	33	3.8	120	2	US-08-819-867-58	Sequence 58, Appli
99	33	3.8	120	4	US-08-464-011B-44	Sequence 44, Appli
100	33	3.8	140	1	US-08-628-417-5	Sequence 5, Appli

101	33	3.8	144	1	US-08-702-344-26	Sequence 26, Appl	174	33	3.8	958	2	US-08-757-046A-5	Sequence 5, Appl1
102	33	3.8	216	1	US-08-666-878A-34	Sequence 34, Appl	175	33	3.8	958	3	US-09-447-208-5	Sequence 5, Appl1
103	33	3.8	216	4	US-09-175-928-34	Sequence 34, Appl	176	33	3.8	958	4	US-09-135-988-5	Sequence 5, Appl1
104	33	3.8	222	4	US-08-481-180-15	Sequence 15, Appl	177	33	3.8	958	4	US-09-277-716-5	Sequence 5, Appl1
105	33	3.8	222	5	PCT-US93-00869-15	Sequence 15, Appl	178	33	3.8	958	4	US-08-597-274A-5	Sequence 5, Appl1
106	33	3.8	223	4	US-09-792-594-11	Sequence 11, Appl	179	33	3.8	958	4	US-08-908-909-5	Sequence 5, Appl1
107	33	3.8	240	1	US-08-628-417-6	Sequence 6, Appl1	180	33	3.8	958	4	US-09-609-161B-5	Sequence 5, Appl1
108	33	3.8	255	4	US-09-480-921B-26	Sequence 26, Appl	181	33	3.8	958	4	US-08-990-103-5	Sequence 5, Appl1
109	33	3.8	340	1	US-08-171-385-27	Sequence 27, Appl	182	33	3.8	966	2	US-08-514-014-7	Sequence 7, Appl1
110	33	3.8	340	1	US-08-361-441B-27	Sequence 27, Appl	183	33	3.8	966	2	US-08-833-823-7	Sequence 7, Appl1
111	33	3.8	347	1	US-08-104-072B-2	Sequence 2, Appl1	184	33	3.8	980	4	US-08-504-459-9	Sequence 9, Appl1
112	33	3.8	350	1	US-08-171-385-14	Sequence 14, Appl	185	33	3.8	980	4	US-09-452-239-21	Sequence 21, Appl
113	33	3.8	350	1	US-08-451-501-11	Sequence 11, Appl	186	33	3.8	991	3	US-08-924-747-25	Sequence 25, Appl
114	33	3.8	350	1	US-08-361-441B-14	Sequence 14, Appl	187	33	3.8	991	4	US-09-247-737B-25	Sequence 25, Appl
115	33	3.8	350	5	PCT-US95-06761-11	Sequence 11, Appl	188	33	3.8	991	4	US-09-296-715-25	Sequence 25, Appl
116	33	3.8	350	5	US-08-465-380-39	Sequence 39, Appl	189	33	3.8	1035	2	US-08-580-545B-3	Sequence 3, Appl1
117	33	3.8	361	2	US-08-465-380-39	Sequence 39, Appl	190	33	3.8	1035	4	US-09-662-653A-3	Sequence 3, Appl1
118	33	3.8	361	2	US-08-486-397-39	Sequence 39, Appl	191	33	3.8	1035	4	US-08-867-484A-1	Sequence 1, Appl1
119	33	3.8	361	2	US-08-486-399-39	Sequence 39, Appl	192	33	3.8	1037	2	US-08-824-405-1	Sequence 1, Appl1
120	33	3.8	361	2	US-08-461-965-39	Sequence 39, Appl	193	33	3.8	1069	4	US-09-372-422A-7	Sequence 7, Appl1
121	33	3.8	361	2	US-08-634-641-39	Sequence 39, Appl	194	33	3.8	1074	4	US-09-248-335-67	Sequence 67, Appl1
122	33	3.8	361	3	US-09-248-471-39	Sequence 39, Appl	195	33	3.8	1075	4	US-08-400-006B-6	Sequence 6, Appl1
123	33	3.8	361	3	US-09-248-472-39	Sequence 39, Appl	196	33	3.8	1100	4	US-09-248-335-53	Sequence 53, Appl
124	33	3.8	361	3	US-09-248-451-39	Sequence 39, Appl	197	33	3.8	1100	4	US-07-861-458C-4	Sequence 4, Appl1
125	33	3.8	361	3	US-08-809-455-39	Sequence 39, Appl	198	33	3.8	1106	4	US-09-362-318-1	Sequence 1, Appl1
126	33	3.8	361	3	US-09-249-461-39	Sequence 39, Appl	199	33	3.8	1114	4	US-09-152-060-41	Sequence 41, Appl
127	33	3.8	361	3	US-09-249-448-39	Sequence 39, Appl	200	33	3.8	1117	4	US-09-247-373B-33	Sequence 33, Appl
128	33	3.8	375	3	US-08-946-026-23	Sequence 23, Appl	201	33	3.8	1123	4	US-09-152-060-15	Sequence 15, Appl
129	33	3.8	399	1	US-07-885-970A-13	Sequence 13, Appl	202	33	3.8	1153	4	US-09-149-476-119	Sequence 119, Appl
130	33	3.8	399	1	US-08-528-687A-13	Sequence 13, Appl	203	33	3.8	1159	4	US-09-410-464-14	Sequence 14, Appl
131	33	3.8	399	1	US-08-510-797-12	Sequence 12, Appl	204	33	3.8	1166	5	PCT-US96-12129B-1	Sequence 9, Appl1
132	33	3.8	399	1	US-08-298-829-13	Sequence 13, Appl	205	33	3.8	1172	1	US-07-945-288-9	Sequence 9, Appl1
133	33	3.8	399	2	US-08-781-035-12	Sequence 12, Appl	206	33	3.8	1172	1	US-08-462-831-9	Sequence 9, Appl1
134	33	3.8	427	4	US-09-091-097-41	Sequence 41, Appl	207	33	3.8	1172	1	US-08-461-809-9	Sequence 9, Appl1
135	33	3.8	458	1	US-08-524-757-1	Sequence 1, Appl1	208	33	3.8	1172	1	US-08-461-441-9	Sequence 9, Appl1
136	33	3.8	488	4	US-09-328-111-1746	Sequence 16, Appl	209	33	3.8	1172	5	PCT-US93-08518-9	Sequence 9, Appl1
137	33	3.8	506	4	US-09-328-111-1746	Sequence 16, Appl	210	33	3.8	1172	5	US-08-872-437-1	Sequence 1, Appl1
138	33	3.8	522	2	US-08-909-965C-16	Sequence 506, Appl	211	33	3.8	1174	3	US-08-651-136C-11	Sequence 11, Appl
139	33	3.8	545	4	US-09-328-111-506	Sequence 20, Appl	212	33	3.8	1174	4	US-09-229-911A-11	Sequence 11, Appl
140	33	3.8	568	1	US-08-583-257-20	Sequence 20, Appl	213	33	3.8	1181	1	US-08-181-271A-27	Sequence 27, Appl
141	33	3.8	568	2	US-08-583-298-08	Sequence 20, Appl	214	33	3.8	1181	1	US-08-449-315-27	Sequence 27, Appl
142	33	3.8	577	4	US-09-149-476-313	Sequence 313, Appl	215	33	3.8	1181	1	US-08-449-315-27	Sequence 27, Appl
143	33	3.8	593	4	US-09-383-982-262	Sequence 262, Appl	216	33	3.8	1181	1	US-08-449-315-27	Sequence 27, Appl
144	33	3.8	646	4	US-09-288-143-42	Sequence 12, Appl	217	33	3.8	1181	1	US-08-456-263A-27	Sequence 27, Appl
145	33	3.8	654	5	PCT-US95-06406A-11	Sequence 11, Appl	218	33	3.8	1181	1	US-08-455-416-27	Sequence 27, Appl
146	33	3.8	704	4	US-09-475-316A-26	Sequence 26, Appl	219	33	3.8	1181	1	US-08-455-416-27	Sequence 27, Appl
147	33	3.8	724	4	US-09-020-956-38	Sequence 38, Appl	220	33	3.8	1181	1	US-08-455-416-27	Sequence 27, Appl
148	33	3.8	724	4	US-09-030-607-38	Sequence 38, Appl	221	33	3.8	1181	1	US-08-457-364-27	Sequence 27, Appl
149	33	3.8	724	4	US-09-605-785-38	Sequence 38, Appl	222	33	3.8	1181	2	US-08-456-240-27	Sequence 27, Appl
150	33	3.8	724	4	US-09-439-313-38	Sequence 38, Appl	223	33	3.8	1181	2	US-08-456-240-27	Sequence 27, Appl
151	33	3.8	724	4	US-09-353-616A-38	Sequence 38, Appl	224	33	3.8	1181	2	US-08-456-240-27	Sequence 27, Appl
152	33	3.8	724	4	US-09-232-149A-38	Sequence 38, Appl	225	33	3.8	1181	2	US-08-455-736-27	Sequence 27, Appl
153	33	3.8	728	4	US-09-091-097-5	Sequence 8, Appl1	226	33	3.8	1181	2	US-08-971-217-27	Sequence 27, Appl
154	33	3.8	740	2	US-08-713-000-8	Sequence 8, Appl1	227	33	3.8	1181	4	US-09-350-600-27	Sequence 27, Appl
155	33	3.8	740	2	US-08-975-316-8	Sequence 8, Appl1	228	33	3.8	1192	4	US-09-318-448-18	Sequence 18, Appl
156	33	3.8	740	4	US-09-211-710-8	Sequence 8, Appl1	229	33	3.8	1198	3	US-09-248-335-27	Sequence 27, Appl
157	33	3.8	740	4	US-09-615-192A-8	Sequence 8, Appl1	230	33	3.8	1212	4	US-09-182-145-34	Sequence 34, Appl
158	33	3.8	741	2	US-08-975-316-58	Sequence 58, Appl	231	33	3.8	1212	4	US-09-182-145-35	Sequence 35, Appl
159	33	3.8	741	2	US-09-615-192A-58	Sequence 58, Appl	232	33	3.8	1223	3	US-09-154-874-4	Sequence 4, Appl1
160	33	3.8	788	4	US-09-615-192A-96	Sequence 96, Appl	233	33	3.8	1229	1	US-08-440-861-1	Sequence 1, Appl1
161	33	3.8	790	4	US-09-363-970-4	Sequence 4, Appl1	234	33	3.8	1229	1	US-08-433-854-1	Sequence 1, Appl1
162	33	3.8	812	4	US-09-091-097-7	Sequence 7, Appl1	235	33	3.8	1229	2	US-08-174-745A-1	Sequence 1, Appl1
163	33	3.8	826	4	US-09-227-357-102	Sequence 102, Appl	236	33	3.8	1229	2	US-08-195-947-1	Sequence 1, Appl1
164	33	3.8	827	5	PCT-US95-06406A-20	Sequence 20, Appl	237	33	3.8	1229	2	US-08-433-885-1	Sequence 1, Appl1
165	33	3.8	836	4	US-09-353-990-7	Sequence 7, Appl1	238	33	3.8	1229	2	US-08-433-885B-1	Sequence 1, Appl1
166	33	3.8	851	4	US-09-443-184-39	Sequence 39, Appl	239	33	3.8	1229	4	US-08-410-614-1	Sequence 1, Appl1
167	33	3.8	857	3	US-08-460-040-1	Sequence 1, Appl1	240	33	3.8	1242	4	US-08-413-974-1	Sequence 1, Appl1
168	33	3.8	911	2	US-08-924-759-9	Sequence 9, Appl1	241	33	3.8	1242	4	US-08-434-418-1	Sequence 1, Appl1
169	33	3.8	911	2	US-09-248-335-9	Sequence 9, Appl1	242	33	3.8	1242	4	US-08-433-888-1	Sequence 1, Appl1
170	33	3.8	936	2	US-08-557-309B-21	Sequence 21, Appl	243	33	3.8	1242	4	US-08-174-739-1	Sequence 1, Appl1
171	33	3.8	936	3	US-08-834-306-21	Sequence 21, Appl	244	33	3.8	1242	4	US-08-434-256-1	Sequence 1, Appl1
172	33	3.8	936	4	US-08-993-674A-21	Sequence 21, Appl	245	33	3.8	1265	4	US-08-991-789A-169	Sequence 169, Appl
173	33	3.8	936	4	US-09-256-976-21	Sequence 21, Appl	246	33	3.8	1265	4	US-09-062-451-169	Sequence 169, Appl

247	33	3.8	1265	4	US-09-598-326-169	Sequence 169, App	320	33	3.8	1728	3	US-08-985-950-7	Sequence 7, Appli
248	33	3.8	1273	1	US-08-507-431-1	Sequence 1, Appli	321	33	3.8	1737	1	US-08-202-056-4	Sequence 4, Appli
249	33	3.8	1273	2	US-08-902-655A-1	Sequence 1, Appli	322	33	3.8	1737	1	US-08-076-033A-3	Sequence 3, Appli
250	33	3.8	1273	3	US-09-116-622-1	Sequence 1, Appli	323	33	3.8	1737	1	US-08-701-265-3	Sequence 3, Appli
251	33	3.8	1273	4	US-09-219-277-1	Sequence 1, Appli	324	33	3.8	1737	2	US-08-284-586-3	Sequence 3, Appli
252	33	3.8	1273	4	US-09-599-661-1	Sequence 1, Appli	325	33	3.8	1737	2	US-08-805-478-3	Sequence 3, Appli
253	33	3.8	1279	4	US-09-277-716-31	Sequence 31, Appli	326	33	3.8	1737	2	US-08-802-627A-3	Sequence 3, Appli
254	33	3.8	1279	4	US-09-609-1618-31	Sequence 31, Appli	327	33	3.8	1737	2	US-08-801-238-3	Sequence 3, Appli
255	33	3.8	1283	1	US-08-174-467-19	Sequence 19, Appli	328	33	3.8	1737	2	US-08-801-228-3	Sequence 3, Appli
256	33	3.8	1283	3	US-08-452-071-19	Sequence 11, Appli	329	33	3.8	1737	5	US-09-104-296-3	Sequence 2, Appli
257	33	3.8	1283	4	US-09-282-305-11	Sequence 10, Appli	330	33	3.8	1737	5	PCT-US94-06380-2	Sequence 2, Appli
258	33	3.8	1307	4	US-09-250-609-3	Sequence 3, Appli	331	33	3.8	1738	2	US-08-379-482A-2	Sequence 56, Appli
259	33	3.8	1315	4	US-09-721-822A-10	Sequence 3, Appli	332	33	3.8	1753	4	US-09-149-476-56	Sequence 5, Appli
260	33	3.8	1323	3	US-09-413-068-3	Sequence 3, Appli	333	33	3.8	1753	4	US-08-836-567-3	Sequence 3, Appli
261	33	3.8	1323	4	US-09-403-345A-3	Sequence 3, Appli	334	33	3.8	1758	3	US-08-655-352-10	Sequence 10, Appli
262	33	3.8	1323	4	US-08-265-087-3	Sequence 3, Appli	335	33	3.8	1776	3	US-09-258-016-10	Sequence 10, Appli
263	33	3.8	1364	1	US-08-621-493-3	Sequence 3, Appli	336	33	3.8	1776	4	US-09-257-825B-10	Sequence 1, Appli
264	33	3.8	1364	1	US-08-872-302-3	Sequence 3, Appli	337	33	3.8	1781	4	US-09-499-302A-1	Sequence 1, Appli
265	33	3.8	1364	2	US-08-965-688-3	Sequence 3, Appli	338	33	3.8	1817	4	US-08-473-981A-5	Sequence 5, Appli
266	33	3.8	1364	2	US-09-260-173-3	Sequence 3, Appli	339	33	3.8	1817	2	US-08-474-087-5	Sequence 5, Appli
267	33	3.8	1364	4	US-09-615-182A-352	Sequence 352, App	340	33	3.8	1821	4	US-09-149-476-90	Sequence 90, Appli
268	33	3.8	1391	4	US-07-602-824A-1	Sequence 1, Appli	341	33	3.8	1824	3	US-08-606-505B-1	Sequence 1, Appli
269	33	3.8	1393	1	US-07-983-451-1	Sequence 1, Appli	342	33	3.8	1824	4	US-09-616-990-1	Sequence 1, Appli
270	33	3.8	1393	1	US-08-261-577-6	Sequence 6, Appli	343	33	3.8	1829	3	US-08-688-988-9	Sequence 9, Appli
271	33	3.8	1393	1	US-09-983-451-1	Sequence 1, Appli	344	33	3.8	1835	4	US-09-485-54-1	Sequence 1, Appli
272	33	3.8	1400	4	US-09-018-584A-35	Sequence 35, Appli	345	33	3.8	1835	4	US-09-286-528-19	Sequence 19, Appli
273	33	3.8	1400	4	US-09-245-281-40	Sequence 40, Appli	346	33	3.8	1859	4	US-08-658-883B-1	Sequence 26, Appli
274	33	3.8	1400	4	US-09-207-359B-40	Sequence 40, Appli	347	33	3.8	1868	4	US-09-676-610B-66	Sequence 1, Appli
275	33	3.8	1420	2	US-08-909-965C-3	Sequence 3, Appli	348	33	3.8	1868	4	US-09-801-052-1	Sequence 1, Appli
276	33	3.8	1430	2	US-09-157-864-3	Sequence 3, Appli	349	33	3.8	1875	5	PCT-US96-10618-10	Sequence 1, Appli
277	33	3.8	1454	4	US-09-372-422A-19	Sequence 19, Appli	350	33	3.8	1883	4	US-09-149-476-170	Sequence 170, App
278	33	3.8	1461	4	US-08-723-126A-4	Sequence 4, Appli	351	33	3.8	1910	2	US-09-009-438-1	Sequence 1, Appli
279	33	3.8	1461	5	PCT-US95-04258-4	Sequence 1, Appli	352	33	3.8	1910	4	US-09-207-493-1	Sequence 1, Appli
280	33	3.8	1463	4	US-09-426-557-1	Sequence 64, Appli	353	33	3.8	1951	4	US-09-465-558-35	Sequence 35, Appli
281	33	3.8	1474	4	US-08-821-994-64	Sequence 87, Appli	354	33	3.8	1964	3	US-08-468-856B-7	Sequence 7, Appli
282	33	3.8	1481	2	US-08-474-379C-87	Sequence 1, Appli	355	33	3.8	1964	4	US-08-468-859A-7	Sequence 68, Appli
283	33	3.8	1481	2	US-09-262-749-1	Sequence 1, Appli	356	33	3.8	1984	4	US-09-232-197-6	Sequence 68, Appli
284	33	3.8	1485	4	US-09-372-422A-39	Sequence 39, Appli	357	33	3.8	1984	4	US-09-232-197-6	Sequence 68, Appli
285	33	3.8	1490	2	US-08-553-367A-5	Sequence 5, Appli	358	33	3.8	1988	4	US-09-332-201-6	Sequence 26, Appli
286	33	3.8	1490	4	US-09-295-306-5	Sequence 5, Appli	359	33	3.8	2000	1	US-08-351-473B-1	Sequence 1, Appli
287	33	3.8	1490	4	US-09-734-719-5	Sequence 5, Appli	360	33	3.8	2000	4	US-09-605-785-374	Sequence 374, App
288	33	3.8	1502	2	US-08-651-940-1	Sequence 1, Appli	361	33	3.8	2000	4	US-09-439-613-374	Sequence 374, App
289	33	3.8	1505	2	US-09-295-029-1	Sequence 13, Appli	362	33	3.8	2000	4	US-09-352-616A-374	Sequence 1, Appli
290	33	3.8	1505	2	US-08-909-965C-13	Sequence 1, Appli	363	33	3.8	2002	4	US-09-819-993-1	Sequence 1, Appli
291	33	3.8	1508	4	US-09-039-046-1	Sequence 1, Appli	364	33	3.8	2026	2	US-08-993-228-3	Sequence 3, Appli
292	33	3.8	1522	4	US-09-413-574-1	Sequence 31, App	365	33	3.8	2055	4	US-09-370-473-5	Sequence 3, Appli
293	33	3.8	1537	4	US-09-149-476-311	Sequence 7, Appli	366	33	3.8	2055	4	US-09-332-191-6	Sequence 6, Appli
294	33	3.8	1540	3	US-08-977-001-2	Sequence 1, Appli	367	33	3.8	2057	4	US-09-332-197-6	Sequence 6, Appli
295	33	3.8	1551	4	US-09-461-474-7	Sequence 10, Appli	368	33	3.8	2087	4	US-09-332-197-6	Sequence 6, Appli
296	33	3.8	1553	4	US-09-022-669-1	Sequence 1, Appli	369	33	3.8	2087	4	US-09-332-197-6	Sequence 6, Appli
297	33	3.8	1553	4	US-09-280-116-10	Sequence 1, Appli	370	33	3.8	2087	4	US-09-332-197-6	Sequence 6, Appli
298	33	3.8	1553	4	US-09-603-567-1	Sequence 7, Appli	371	33	3.8	2172	2	US-08-417-174-26	Sequence 26, Appli
299	33	3.8	1556	4	US-09-334-601-7	Sequence 1, Appli	372	33	3.8	2172	2	US-08-231-565A-26	Sequence 26, Appli
300	33	3.8	1578	4	US-09-416-050A-1	Sequence 1, Appli	373	33	3.8	2172	2	US-09-007-961-26	Sequence 26, Appli
301	33	3.8	1578	4	US-09-664-800-1	Sequence 1, Appli	374	33	3.8	2172	4	US-09-267-439-26	Sequence 26, Appli
302	33	3.8	1578	4	US-09-665-309-1	Sequence 1, Appli	375	33	3.8	2179	4	US-09-370-838-152	Sequence 152, App
303	33	3.8	1578	4	US-09-661-569-1	Sequence 1, Appli	376	33	3.8	2186	4	US-09-149-476-163	Sequence 163, App
304	33	3.8	1582	3	US-08-545-196B-10	Sequence 12, Appli	377	33	3.8	2202	2	US-09-388-743-1	Sequence 1, Appli
305	33	3.8	1582	3	US-08-545-196B-12	Sequence 9, Appli	378	33	3.8	2218	2	US-08-985-090-4	Sequence 31, Appli
306	33	3.8	1604	1	US-08-665-966-9	Sequence 9, Appli	379	33	3.8	2222	4	US-09-165-543-31	Sequence 1, Appli
307	33	3.8	1604	1	US-09-041-780-9	Sequence 3, Appli	380	33	3.8	2222	4	US-09-197-380-1	Sequence 24, Appli
308	33	3.8	1609	4	US-08-926-808C-2	Sequence 2, Appli	381	33	3.8	2230	4	US-08-378-313-24	Sequence 1, Appli
309	33	3.8	1637	4	US-08-852-824-3	Sequence 3, Appli	382	33	3.8	2237	1	US-08-467-135B-1	Sequence 1, Appli
310	33	3.8	1639	4	US-09-362-473-5	Sequence 5, Appli	383	33	3.8	2237	2	US-08-915-972A-1	Sequence 1, Appli
311	33	3.8	1646	4	US-08-991-789A-290	Sequence 290, App	384	33	3.8	2237	2	US-09-177-909-1	Sequence 1, Appli
312	33	3.8	1646	4	US-09-062-451-290	Sequence 290, App	385	33	3.8	2239	4	US-09-196-190-1	Sequence 1, Appli
313	33	3.8	1651	4	US-09-465-558-49	Sequence 49, Appli	386	33	3.8	2255	4	US-08-871-572B-3	Sequence 3, Appli
314	33	3.8	1659	1	US-08-333-358-7	Sequence 7, Appli	387	33	3.8	2274	4	US-09-388-743-17	Sequence 17, Appli
315	33	3.8	1659	1	US-08-463-694-7	Sequence 7, Appli	388	33	3.8	2280	4	US-09-366-408-3	Sequence 3, Appli
316	33	3.8	1659	1	US-08-694-501-7	Sequence 7, Appli	389	33	3.8	2301	4	US-09-232-191-8	Sequence 8, Appli
317	33	3.8	1659	1	US-08-548-509-4	Sequence 4, Appli	390	33	3.8	2301	4	US-09-232-200-8	Sequence 8, Appli
318	33	3.8	1700	2	US-08-897-340-4	Sequence 4, Appli	391	33	3.8	2301	4	US-09-232-197-8	Sequence 8, Appli
319	33	3.8	1700	3	US-09-252-329-4	Sequence 4, Appli	392	33	3.8	2301	4	US-09-232-201-8	Sequence 8, Appli

393	3.8	2320	4	US-09-813-817-1	Sequence 1, Appl	466	33	3.8	3410	4	US-09-232-149A-110	Sequence 110, App
394	3.8	2320	4	US-09-978-197-1	Sequence 1, Appl	467	33	3.8	3437	4	US-08-860-339-17	Sequence 17, Appl
395	3.8	2338	4	US-09-232-200-66	Sequence 66, Appl	468	33	3.8	3527	2	US-08-909-965C-7	Sequence 7, Appl
396	3.8	2338	4	US-09-233-197-66	Sequence 66, Appl	469	33	3.8	3581	2	US-08-738-349-1	Sequence 1, Appl
397	3.8	2338	4	US-09-233-201-66	Sequence 20, Appl	470	33	3.8	3602	4	US-09-402-929-1	Sequence 1, Appl
398	3.8	2338	4	US-08-803-805D-20	Sequence 20, Appl	471	33	3.8	3680	4	US-09-647-390-15	Sequence 15, Appl
399	3.8	2399	4	US-09-385-801-1	Sequence 1, Appl	472	33	3.8	3709	4	US-09-541-782-7	Sequence 7, Appl
400	3.8	2405	1	US-08-454-097-30	Sequence 30, Appl	473	33	3.8	3709	4	US-09-723-820-7	Sequence 7, Appl
401	3.8	2405	1	US-08-185-359-30	Sequence 30, Appl	474	33	3.8	3825	4	US-09-208-742-3	Sequence 3, Appl
402	3.8	2409	4	US-09-149-476-228	Sequence 228, App	475	33	3.8	4482	2	US-08-567-508C-1	Sequence 1, Appl
403	3.8	2424	4	US-09-293-322C-4	Sequence 1, Appl	476	33	3.8	4482	3	US-09-196-480-1	Sequence 1, Appl
404	3.8	2484	4	US-09-293-322C-4	Sequence 1, Appl	477	33	3.8	4655	3	US-08-948-378A-7	Sequence 7, Appl
405	3.8	2502	3	US-09-233-332-1	Sequence 1, Appl	478	33	3.8	4655	4	US-09-169-425C-7	Sequence 7, Appl
406	3.8	2550	6	US-08-569-749-1	Sequence 1, Appl	479	33	3.8	4931	4	US-08-726-320-2	Sequence 2, Appl
407	3.8	2589	4	US-08-569-749-1	Sequence 1, Appl	480	33	3.8	4931	4	US-09-108-718-2	Sequence 2, Appl
408	3.8	2589	5	PCT-US96-12860-1	Sequence 1, Appl	481	33	3.8	5503	2	US-08-726-012B-1	Sequence 1, Appl
409	3.8	2601	5	US-08-569-749-3	Sequence 3, Appl	482	33	3.8	6200	4	US-09-439-923-1	Sequence 1, Appl
410	3.8	2601	5	PCT-US96-12860-3	Sequence 3, Appl	483	33	3.8	6638	2	US-08-070-301-2	Sequence 2, Appl
411	3.8	2619	4	US-08-983-502-17	Sequence 17, Appl	484	33	3.8	7941	4	US-09-816-703A-1	Sequence 1, Appl
412	3.8	2619	5	PCT-US96-10521-17	Sequence 17, Appl	485	33	3.8	55827	4	US-09-813-133A-3	Sequence 3, Appl
413	3.8	2625	4	US-09-245-041-18	Sequence 18, Appl	486	33	3.8	59065	4	US-09-813-817-3	Sequence 3, Appl
414	3.8	2628	2	US-08-696-944-1	Sequence 1, Appl	487	33	3.8	59065	4	US-09-978-197-3	Sequence 3, Appl
415	3.8	2634	2	US-08-818-514-1	Sequence 1, Appl	488	33	3.8	80246	4	US-09-078-294-4	Sequence 4, Appl
416	3.8	2634	2	US-08-818-514-2	Sequence 2, Appl	489	33	3.7	36	4	US-09-462-569E-4	Sequence 4, Appl
417	3.8	2634	2	US-09-115-934A-1	Sequence 2, Appl	490	33	3.7	37	1	US-08-113-646A-44	Sequence 4, Appl
418	3.8	2634	3	US-09-115-934A-2	Sequence 2, Appl	491	33	3.7	40	2	US-08-771-624B-1	Sequence 1, Appl
419	3.8	2658	2	US-08-593-383-3	Sequence 3, Appl	492	33	3.7	40	2	US-08-440-209-4	Sequence 4, Appl
420	3.8	2671	6	US-09-233-200-44	Sequence 4, Appl	493	33	3.7	40	3	US-08-439-996-4	Sequence 7, Appl
421	3.8	2710	4	US-09-233-200-44	Sequence 4, Appl	494	33	3.7	40	3	US-09-732-067-7	Sequence 4, Appl
422	3.8	2710	4	US-09-233-200-70	Sequence 4, Appl	495	33	3.7	44	1	US-08-113-646A-40	Sequence 1, Appl
423	3.8	2710	4	US-09-233-197-44	Sequence 44, Appl	496	33	3.7	50	1	US-08-420-443-1	Sequence 1, Appl
424	3.8	2710	4	US-09-233-197-70	Sequence 70, Appl	497	33	3.7	52	4	US-09-415-784-23	Sequence 23, Appl
425	3.8	2710	4	US-09-233-201-44	Sequence 44, Appl	498	33	3.7	52	4	US-08-944-465-23	Sequence 23, Appl
426	3.8	2710	4	US-09-233-201-70	Sequence 70, Appl	499	33	3.7	52	4	US-09-415-868-23	Sequence 23, Appl
427	3.8	2719	3	US-08-706-216-1	Sequence 1, Appl	500	33	3.7	52	4	US-09-415-900-23	Sequence 23, Appl
428	3.8	2790	3	US-08-800-291B-1	Sequence 1, Appl	501	33	3.7	57	4	US-09-199-737-59	Sequence 59, Appl
429	3.8	2836	3	US-08-747-221B-24	Sequence 24, Appl	502	33	3.7	57	4	US-09-058-333A-59	Sequence 59, Appl
430	3.8	2836	4	US-09-005-051-24	Sequence 24, Appl	503	33	3.7	58	2	US-08-417-174-42	Sequence 42, Appl
431	3.8	2836	4	US-09-005-051-24	Sequence 24, Appl	504	33	3.7	58	2	US-08-231-655A-40	Sequence 40, Appl
432	3.8	2837	2	US-08-993-228-11	Sequence 11, Appl	505	33	3.7	58	2	US-09-007-961-40	Sequence 40, Appl
433	3.8	2887	5	PCT-US96-10521-14	Sequence 14, Appl	506	33	3.7	58	2	US-09-267-439-42	Sequence 42, Appl
434	3.8	2887	5	US-08-983-502-14	Sequence 14, Appl	507	33	3.7	60	1	US-08-241-645B-11	Sequence 11, Appl
435	3.8	2887	5	US-08-927-394-1	Sequence 1, Appl	508	33	3.7	60	1	US-09-457-959-8	Sequence 8, Appl
436	3.8	2897	4	US-09-233-200-52	Sequence 52, Appl	509	33	3.7	61	4	US-09-415-785A-32	Sequence 32, Appl
437	3.8	2897	4	US-09-233-200-52	Sequence 52, Appl	510	33	3.7	65	4	US-09-415-785A-32	Sequence 32, Appl
438	3.8	2907	4	US-09-233-201-52	Sequence 52, Appl	511	33	3.7	65	4	US-08-944-465-32	Sequence 32, Appl
439	3.8	2907	4	US-09-233-200-52	Sequence 52, Appl	512	33	3.7	65	4	US-08-944-465-32	Sequence 32, Appl
440	3.8	2917	4	US-09-232-197-26	Sequence 26, Appl	513	33	3.7	65	4	US-09-415-868-32	Sequence 32, Appl
441	3.8	2917	4	US-09-232-197-26	Sequence 26, Appl	514	33	3.7	65	4	US-09-415-868-32	Sequence 32, Appl
442	3.8	2940	2	US-08-593-383-1	Sequence 1, Appl	515	33	3.7	65	4	US-08-702-344-7	Sequence 7, Appl
443	3.8	2940	6	US-08-593-383-1	Sequence 1, Appl	516	33	3.7	69	1	US-08-702-344-7	Sequence 7, Appl
444	3.8	2940	6	US-08-593-383-1	Sequence 1, Appl	517	33	3.7	69	1	US-08-702-344-7	Sequence 7, Appl
445	3.8	2992	4	US-09-334-601-11	Sequence 11, Appl	518	33	3.7	70	2	US-08-776-944-13	Sequence 13, Appl
446	3.8	3073	1	US-07-688-352C-11	Sequence 11, Appl	519	33	3.7	75	2	US-08-776-944-13	Sequence 13, Appl
447	3.8	3073	2	US-08-474-379C-11	Sequence 31, Appl	520	33	3.7	80	1	US-07-920-281C-25	Sequence 25, Appl
448	3.8	3073	2	US-08-146-249A-11	Sequence 31, Appl	521	33	3.7	80	1	US-08-466-277-25	Sequence 25, Appl
449	3.8	3073	3	US-08-206-188B-31	Sequence 31, Appl	522	33	3.7	84	1	US-08-664-596B-3	Sequence 3, Appl
450	3.8	3073	5	PCT-US91-02714-30	Sequence 30, Appl	523	33	3.7	84	1	US-08-738-367-3	Sequence 3, Appl
451	3.8	3131	1	US-07-688-352C-21	Sequence 21, Appl	524	33	3.7	85	1	US-08-120-827-97	Sequence 97, Appl
452	3.8	3131	3	US-09-146-249A-11	Sequence 21, Appl	525	33	3.7	85	1	US-08-478-675-97	Sequence 97, Appl
453	3.8	3131	3	US-08-206-188B-31	Sequence 21, Appl	526	33	3.7	90	1	US-08-677-944-1	Sequence 1, Appl
454	3.8	3131	5	PCT-US91-02714-70	Sequence 20, Appl	527	33	3.7	90	1	US-08-677-944-1	Sequence 1, Appl
455	3.8	3159	4	US-09-437-054A-7	Sequence 7, Appl	528	33	3.7	90	4	US-09-065-058-16	Sequence 16, Appl
456	3.8	3214	1	US-08-484-105-17	Sequence 17, Appl	529	33	3.7	90	4	US-09-254-048A-1	Sequence 1, Appl
457	3.8	3214	1	US-08-484-105-17	Sequence 17, Appl	530	33	3.7	90	4	US-09-921-303-1	Sequence 1, Appl
458	3.8	3244	3	US-09-165-543-4	Sequence 4, Appl	531	33	3.7	91	4	US-09-404-879A-201	Sequence 201, App
459	3.8	3244	3	US-09-165-543-4	Sequence 4, Appl	532	33	3.7	94	4	US-09-404-879A-261	Sequence 261, App
460	3.8	3375	4	US-09-511-625B-67	Sequence 67, Appl	533	33	3.7	98	1	US-08-088-658A-42	Sequence 42, Appl
461	3.8	3410	4	US-09-020-956-110	Sequence 110, App	534	33	3.7	98	2	US-08-471-907A-42	Sequence 42, Appl
462	3.8	3410	4	US-09-030-607-110	Sequence 110, App	535	33	3.7	100	4	US-08-991-789A-30	Sequence 30, Appl
463	3.8	3410	4	US-09-605-785-110	Sequence 110, App	536	33	3.7	100	4	US-09-062-451-30	Sequence 30, Appl
464	3.8	3410	4	US-09-439-313-110	Sequence 110, App	537	33	3.7	100	4	US-09-062-451-30	Sequence 30, Appl
465	3.8	3410	4	US-09-352-616A-110	Sequence 110, App	538	33	3.7	101	4	US-09-404-879A-293	Sequence 293, App



539	32	3.7	105	2	US-08-735-381-2	Sequence 2, Appl1	C 612	32	3.7	470	4	US-09-439-313-102	Sequence 102, App
540	32	3.7	105	3	US-09-183-619-1	Sequence 1, Appl1	C 613	32	3.7	470	4	US-09-352-616A-102	Sequence 102, App
541	32	3.7	105	3	US-09-201-674-2	Sequence 2, Appl1	C 614	32	3.7	470	4	US-09-232-149A-102	Sequence 102, App
542	32	3.7	111	4	US-09-297-535-23	Sequence 23, Appl1	C 615	32	3.7	474	4	US-08-516-859A-97	Sequence 97, Appl1
543	32	3.7	117	1	US-08-702-344-3	Sequence 3, Appl1	C 616	32	3.7	474	4	US-09-586-472-97	Sequence 97, Appl1
544	32	3.7	121	4	US-09-297-535-20	Sequence 20, Appl1	C 617	32	3.7	474	4	US-09-528-706-97	Sequence 29, Appl1
545	32	3.7	130	6	5198345-15	Patent No. 5198345	C 618	32	3.7	502	4	US-09-186-276B-29	Sequence 29, Appl1
546	32	3.7	130	6	5198345-15	Patent No. 5198345	C 619	32	3.7	502	4	US-08-842-445-29	Sequence 29, Appl1
547	32	3.7	184	4	US-09-480-921B-13	Sequence 13, Appl1	C 620	32	3.7	502	4	US-09-166-188B-29	Sequence 29, Appl1
548	32	3.7	198	1	US-08-330-108-16	Sequence 16, Appl1	C 621	32	3.7	509	4	US-09-030-607-202	Sequence 202, App
549	32	3.7	198	5	PCT-US92-10087-16	Sequence 16, Appl1	C 622	32	3.7	509	4	US-09-605-785-202	Sequence 202, App
550	32	3.7	208	1	US-08-686-878A-37	Sequence 37, Appl1	C 623	32	3.7	509	4	US-09-439-313-202	Sequence 202, App
551	32	3.7	208	2	US-08-967-101-98	Sequence 98, Appl1	C 624	32	3.7	509	4	US-09-352-616A-202	Sequence 202, App
552	32	3.7	208	2	US-08-592-541-98	Sequence 98, Appl1	C 625	32	3.7	515	3	US-08-589-028-1	Sequence 1, Appl1
553	32	3.7	208	3	US-09-124-698-98	Sequence 98, Appl1	C 626	32	3.7	515	3	US-08-784-582-1	Sequence 1, Appl1
554	32	3.7	208	4	US-09-127-480-98	Sequence 98, Appl1	C 627	32	3.7	515	3	US-08-785-271-1	Sequence 12, Appl1
555	32	3.7	208	4	US-08-496-841C-98	Sequence 98, Appl1	C 628	32	3.7	528	1	US-09-105-542A-12	Sequence 1, Appl1
556	32	3.7	208	4	US-09-175-928-37	Sequence 37, Appl1	C 630	32	3.7	536	2	US-08-911-020-1	Sequence 1, Appl1
557	32	3.7	208	4	US-09-124-523-98	Sequence 98, Appl1	C 631	32	3.7	536	2	US-08-911-020-1	Sequence 1, Appl1
558	32	3.7	227	4	US-08-897-126-28	Sequence 28, Appl1	C 632	32	3.7	556	4	US-09-370-838-242	Sequence 242, App
559	32	3.7	227	4	US-08-897-126-28	Sequence 28, Appl1	C 633	32	3.7	572	4	US-09-328-111-644	Sequence 644, App
560	32	3.7	227	4	US-09-191-136-29	Sequence 29, Appl1	C 634	32	3.7	572	4	US-09-342-653-5	Sequence 5, Appl1
561	32	3.7	253	2	US-08-520-678A-25	Sequence 25, Appl1	C 635	32	3.7	578	4	US-09-385-982-203	Sequence 757, App
562	32	3.7	253	4	US-08-897-126-25	Sequence 25, Appl1	C 636	32	3.7	578	4	US-09-602-877A-95	Sequence 95, Appl1
563	32	3.7	257	2	US-08-520-678A-24	Sequence 24, Appl1	C 637	32	3.7	581	2	US-08-557-309B-22	Sequence 22, Appl1
564	32	3.7	257	2	US-08-897-126-24	Sequence 24, Appl1	C 638	32	3.7	581	2	US-08-834-306-22	Sequence 22, Appl1
565	32	3.7	260	4	US-08-520-678A-29	Sequence 29, Appl1	C 639	32	3.7	581	3	US-08-993-674A-22	Sequence 22, Appl1
566	32	3.7	260	4	US-08-897-126-29	Sequence 29, Appl1	C 640	32	3.7	581	4	US-09-255-976-22	Sequence 22, Appl1
567	32	3.7	263	4	US-09-091-097-26	Sequence 26, Appl1	C 641	32	3.7	588	4	US-09-385-982-128	Sequence 128, App
568	32	3.7	270	2	US-08-520-678A-30	Sequence 30, Appl1	C 642	32	3.7	588	4	US-08-150-331-45	Sequence 45, Appl1
569	32	3.7	270	2	US-08-897-126-30	Sequence 30, Appl1	C 643	32	3.7	596	1	US-08-150-331-45	Sequence 147, App
570	32	3.7	271	2	US-08-731-272A-29	Sequence 29, Appl1	C 644	32	3.7	599	4	US-09-328-111-147	Sequence 6, Appl1
571	32	3.7	288	2	US-08-648-496-1	Sequence 1, Appl1	C 645	32	3.7	603	2	US-08-924-838-6	Sequence 27, Appl1
572	32	3.7	290	4	US-09-232-575-20	Sequence 20, Appl1	C 646	32	3.7	603	4	US-09-385-982-293	Sequence 293, App
573	32	3.7	291	4	US-09-329-796-13	Sequence 13, Appl1	C 647	32	3.7	611	4	US-09-105-542A-2	Sequence 2, Appl1
574	32	3.7	296	2	US-09-032-684-13	Sequence 13, Appl1	C 648	32	3.7	615	4	US-09-185-982-376	Sequence 376, App
575	32	3.7	298	2	US-08-520-678A-21	Sequence 21, Appl1	C 649	32	3.7	624	4	US-09-385-982-305	Sequence 305, App
576	32	3.7	298	4	US-08-897-126-21	Sequence 21, Appl1	C 650	32	3.7	624	4	US-09-385-982-305	Sequence 305, App
577	32	3.7	308	2	US-08-721-488-4	Sequence 4, Appl1	C 651	32	3.7	630	1	US-08-185-414E-1	Sequence 1, Appl1
578	32	3.7	314	1	US-08-686-878A-25	Sequence 25, Appl1	C 652	32	3.7	635	1	US-08-455-633A-35	Sequence 35, Appl1
579	32	3.7	314	4	US-09-175-928-25	Sequence 25, Appl1	C 653	32	3.7	635	2	US-08-456-460C-35	Sequence 35, Appl1
580	32	3.7	321	4	US-08-520-678A-23	Sequence 23, Appl1	C 654	32	3.7	635	2	PCT-US94-05314-35	Sequence 327, App
581	32	3.7	321	4	US-08-897-126-23	Sequence 23, Appl1	C 655	32	3.7	635	5	US-08-456-460C-35	Sequence 327, App
582	32	3.7	326	4	US-09-629-645A-17	Sequence 17, Appl1	C 656	32	3.7	635	5	US-09-385-982-327	Sequence 27, Appl1
583	32	3.7	327	4	US-09-385-982-544	Sequence 544, App	C 657	32	3.7	635	5	US-08-226-264-27	Sequence 66, Appl1
584	32	3.7	334	2	US-09-033-684-8	Sequence 8, Appl1	C 658	32	3.7	685	4	US-09-227-357-66	Sequence 285, App
585	32	3.7	337	2	US-09-033-684-11	Sequence 11, Appl1	C 659	32	3.7	688	6	5498694-3	Sequence 285, App
586	32	3.7	356	2	US-08-520-678A-22	Sequence 22, Appl1	C 660	32	3.7	700	1	US-07-846-992-1	Sequence 1, Appl1
587	32	3.7	356	4	US-08-897-126-22	Sequence 22, Appl1	C 661	32	3.7	700	1	US-08-469-555-1	Sequence 26, Appl1
588	32	3.7	374	2	US-08-370-156-24	Sequence 24, Appl1	C 662	32	3.7	700	4	US-09-152-060-26	Sequence 6, Appl1
589	32	3.7	374	4	US-09-385-982-494	Sequence 494, App	C 663	32	3.7	703	4	US-09-313-300-6	Sequence 318, App
590	32	3.7	380	1	US-08-126-587C-5	Sequence 5, Appl1	C 664	32	3.7	712	4	US-09-437-054A-15	Sequence 15, Appl1
591	32	3.7	393	4	US-09-605-785-357	Sequence 357, App	C 665	32	3.7	712	4	US-09-437-054A-15	Sequence 15, Appl1
592	32	3.7	393	4	US-09-439-313-357	Sequence 357, App	C 666	32	3.7	712	4	US-09-149-476-318	Sequence 6, Appl1
593	32	3.7	393	4	US-09-352-616A-357	Sequence 357, App	C 667	32	3.7	712	4	US-09-149-476-318	Sequence 6, Appl1
594	32	3.7	393	4	US-09-195-106-21	Sequence 21, Appl1	C 668	32	3.7	732	1	US-07-847-010-12	Sequence 12, Appl1
595	32	3.7	403	4	US-09-387-212-2	Sequence 2, Appl1	C 669	32	3.7	732	1	US-08-530-165-1	Sequence 1, Appl1
596	32	3.7	403	4	US-09-948-802-2	Sequence 2, Appl1	C 670	32	3.7	732	1	US-08-938-675A-1	Sequence 1, Appl1
597	32	3.7	413	4	US-09-227-357-71	Sequence 71, Appl1	C 671	32	3.7	732	3	US-09-149-476-20	Sequence 20, Appl1
598	32	3.7	429	1	US-08-686-878A-8	Sequence 8, Appl1	C 672	32	3.7	732	3	US-09-361-434-8	Sequence 8, Appl1
599	32	3.7	429	4	US-09-175-928-13	Sequence 13, Appl1	C 673	32	3.7	732	3	US-09-361-434-8	Sequence 8, Appl1
600	32	3.7	433	1	US-07-987-272A-13	Sequence 13, Appl1	C 674	32	3.7	732	3	US-09-361-434-8	Sequence 8, Appl1
601	32	3.7	450	2	US-08-592-541-7	Sequence 7, Appl1	C 675	32	3.7	732	3	US-09-361-434-8	Sequence 8, Appl1
602	32	3.7	450	2	US-08-592-541-7	Sequence 7, Appl1	C 676	32	3.7	732	3	US-09-361-434-8	Sequence 8, Appl1
603	32	3.7	450	2	US-09-124-698-7	Sequence 7, Appl1	C 677	32	3.7	732	3	US-09-361-434-8	Sequence 8, Appl1
604	32	3.7	450	4	US-09-127-480-7	Sequence 7, Appl1	C 678	32	3.7	732	3	US-09-361-434-8	Sequence 8, Appl1
605	32	3.7	450	4	US-08-496-841C-7	Sequence 7, Appl1	C 679	32	3.7	732	3	US-09-361-434-8	Sequence 8, Appl1
606	32	3.7	450	4	US-09-124-523-7	Sequence 7, Appl1	C 680	32	3.7	732	3	US-09-361-434-8	Sequence 8, Appl1
607	32	3.7	464	4	US-09-302-769-37	Sequence 37, Appl1	C 681	32	3.7	732	3	US-09-361-434-8	Sequence 8, Appl1
608	32	3.7	464	2	US-08-841-349-18	Sequence 18, Appl1	C 682	32	3.7	732	3	US-09-361-434-8	Sequence 8, Appl1
609	32	3.7	470	4	US-09-020-956-102	Sequence 102, App	C 683	32	3.7	732	3	US-09-361-434-8	Sequence 8, Appl1
610	32	3.7	470	4	US-09-030-607-102	Sequence 102, App	C 684	32	3.7	732	3	US-09-361-434-8	Sequence 8, Appl1
611	32	3.7	470	4	US-09-605-785-102	Sequence 102, App	C 684	32	3.7	732	3	US-09-361-434-8	Sequence 8, Appl1

685	32	3.7	794	4	US-09-385-982-537	Sequence 537, App	758	32	3.7	1013	1	US-08-314-586-30	Sequence 30, App1
686	32	3.7	804	4	US-09-149-476-125	Sequence 125, App	759	32	3.7	1023	1	US-08-252-966B-16	Sequence 16, App1
687	32	3.7	806	4	US-08-955-629C-1	Sequence 1, App1	760	32	3.7	1027	4	US-09-344-529-1	Sequence 1, App1
688	32	3.7	809	4	US-09-149-476-244	Sequence 244, App	761	32	3.7	1027	4	US-09-465-558-57	Sequence 57, App1
689	32	3.7	817	4	US-08-885-469-1	Sequence 1, App1	762	32	3.7	1032	4	US-09-257-179-21	Sequence 21, App1
690	32	3.7	817	4	US-09-625-918-1	Sequence 1, App1	763	32	3.7	1033	4	US-09-173-300-46	Sequence 46, App1
691	32	3.7	826	1	US-08-658-551-3	Sequence 3, App1	764	32	3.7	1037	4	US-09-257-179-16	Sequence 16, App1
692	32	3.7	826	1	US-08-602-228-3	Sequence 3, App1	765	32	3.7	1046	1	US-08-361-467B-4	Sequence 4, App1
693	32	3.7	826	1	US-08-649-414A-3	Sequence 3, App1	766	32	3.7	1046	1	US-08-484-332C-4	Sequence 4, App1
694	32	3.7	826	1	US-08-494-440B-3	Sequence 3, App1	767	32	3.7	1047	1	US-08-403-378B-3	Sequence 3, App1
695	32	3.7	826	2	US-08-533-901B-3	Sequence 3, App1	768	32	3.7	1051	4	US-09-245-041-10	Sequence 10, App1
696	32	3.7	826	2	US-08-839-032A-3	Sequence 3, App1	769	32	3.7	1052	4	US-09-592-891A-8	Sequence 8, App1
697	32	3.7	826	2	US-08-839-031A-3	Sequence 3, App1	770	32	3.7	1056	4	US-08-157-101A-8	Sequence 4, App1
698	32	3.7	826	5	US-09-185-258C-3	Sequence 3, App1	771	32	3.7	1078	4	US-09-452-757B-1	Sequence 41, App1
699	32	3.7	826	5	PCT-US95-1272A-3	Sequence 3, App1	772	32	3.7	1080	4	US-08-225-757B-1	Sequence 1, App1
700	32	3.7	836	2	US-08-658-805-7	Sequence 7, App1	773	32	3.7	1080	2	US-08-722-050-1	Sequence 1, App1
701	32	3.7	857	1	US-08-308-883-1	Sequence 1, App1	774	32	3.7	1092	4	US-09-227-357-35	Sequence 35, App1
702	32	3.7	857	1	US-08-730-163-1	Sequence 1, App1	775	32	3.7	1098	3	US-09-248-335-35	Sequence 35, App1
703	32	3.7	857	4	US-08-256-799-1	Sequence 1, App1	776	32	3.7	1106	3	US-09-361-434-16	Sequence 16, App1
704	32	3.7	859	4	US-08-462-437-1	Sequence 1, App1	777	32	3.7	1106	3	US-09-361-434-16	Sequence 16, App1
705	32	3.7	859	4	US-09-247-373B-47	Sequence 47, App1	778	32	3.7	1106	4	US-09-635-025-16	Sequence 16, App1
706	32	3.7	860	1	US-07-847-010-18	Sequence 18, App1	779	32	3.7	1113	2	US-08-907-492A-3	Sequence 18, App1
707	32	3.7	872	3	US-09-248-335-63	Sequence 63, App1	780	32	3.7	1116	3	US-08-916-443A-16	Sequence 16, App1
708	32	3.7	873	4	US-09-475-316A-20	Sequence 20, App1	781	32	3.7	1128	3	US-09-136-628-1	Sequence 1, App1
709	32	3.7	879	1	US-08-158-662A-1	Sequence 1, App1	782	32	3.7	1128	3	US-09-227-357-40	Sequence 40, App1
710	32	3.7	879	1	US-08-015-203-1	Sequence 1, App1	783	32	3.7	1129	4	US-09-248-335-29	Sequence 29, App1
711	32	3.7	880	1	US-08-616-368A-7	Sequence 7, App1	784	32	3.7	1134	3	US-09-452-239-3	Sequence 3, App1
712	32	3.7	880	3	US-09-054-298-7	Sequence 7, App1	785	32	3.7	1147	1	US-08-665-716-1	Sequence 1, App1
713	32	3.7	880	4	US-09-247-373B-39	Sequence 39, App1	786	32	3.7	1149	4	US-09-227-357-84	Sequence 84, App1
714	32	3.7	880	4	US-08-818-655-7	Sequence 7, App1	787	32	3.7	1153	4	US-09-149-476-41	Sequence 41, App1
715	32	3.7	882	2	US-08-909-965C-9	Sequence 9, App1	788	32	3.7	1159	4	US-09-100-391-5	Sequence 5, App1
716	32	3.7	888	1	US-08-173-510B-99	Sequence 99, App1	789	32	3.7	1169	4	US-09-149-476-310	Sequence 310, App1
717	32	3.7	888	1	US-08-458-218-97	Sequence 97, App1	790	32	3.7	1182	4	US-09-469-242-1	Sequence 1, App1
718	32	3.7	888	2	US-08-450-497-99	Sequence 99, App1	791	32	3.7	1182	4	US-09-372-422A-23	Sequence 23, App1
719	32	3.7	888	3	US-09-188-930-13	Sequence 13, App1	792	32	3.7	1193	4	US-09-465-558-53	Sequence 53, App1
720	32	3.7	893	1	US-08-276-452A-66	Sequence 66, App1	793	32	3.7	1206	4	US-09-443-074A-29	Sequence 29, App1
721	32	3.7	893	2	US-08-798-744-66	Sequence 66, App1	794	32	3.7	1210	4	US-09-149-476-186	Sequence 186, App1
722	32	3.7	907	4	US-09-149-476-195	Sequence 195, App1	795	32	3.7	1212	4	US-09-605-785-225	Sequence 225, App1
723	32	3.7	911	4	US-09-227-357-109	Sequence 109, App1	796	32	3.7	1214	4	US-09-439-313-225	Sequence 225, App1
724	32	3.7	911	4	US-09-364-230-25	Sequence 25, App1	797	32	3.7	1214	4	US-09-352-616A-225	Sequence 225, App1
725	32	3.7	913	4	US-08-651-136C-13	Sequence 13, App1	798	32	3.7	1214	4	US-09-232-149A-225	Sequence 225, App1
726	32	3.7	913	4	US-09-228-911A-13	Sequence 13, App1	799	32	3.7	1215	4	US-08-592-214A-1	Sequence 1, App1
727	32	3.7	917	4	US-09-227-357-43	Sequence 43, App1	800	32	3.7	1215	2	US-08-370-156-26	Sequence 26, App1
728	32	3.7	919	2	US-08-924-759-19	Sequence 19, App1	801	32	3.7	1215	3	US-09-149-976-1	Sequence 1, App1
729	32	3.7	919	3	US-09-248-335-19	Sequence 19, App1	802	32	3.7	1215	3	US-09-605-785-326	Sequence 326, App1
730	32	3.7	921	4	US-09-227-357-30	Sequence 30, App1	803	32	3.7	1215	4	US-09-439-313-326	Sequence 326, App1
731	32	3.7	940	2	US-08-471-717-1	Sequence 1, App1	804	32	3.7	1215	4	US-09-352-616A-336	Sequence 336, App1
732	32	3.7	940	4	US-09-069-023-9	Sequence 9, App1	805	32	3.7	1215	4	US-09-352-616A-336	Sequence 336, App1
733	32	3.7	940	4	US-09-091-097-11	Sequence 11, App1	806	32	3.7	1215	6	5169941-7	Patent No. 5169941
734	32	3.7	943	4	US-09-149-476-35	Sequence 35, App1	807	32	3.7	1215	6	5169941-7	Patent No. 5169941
735	32	3.7	946	3	US-08-916-443A-15	Sequence 15, App1	808	32	3.7	1217	4	US-09-225-244-1	Sequence 1, App1
736	32	3.7	953	1	US-08-197-793-1	Sequence 1, App1	809	32	3.7	1217	4	US-09-417-242-1	Sequence 1, App1
737	32	3.7	953	2	US-08-636-176-1	Sequence 1, App1	810	32	3.7	1225	1	US-08-286-020-1	Sequence 1, App1
738	32	3.7	953	5	PCT-US95-01618-1	Sequence 1, App1	811	32	3.7	1225	1	US-08-603-919-1	Sequence 1, App1
739	32	3.7	960	3	US-09-248-335-57	Sequence 57, App1	812	32	3.7	1243	1	US-08-314-309A-24	Sequence 24, App1
740	32	3.7	964	1	US-08-448-606-5	Sequence 5, App1	813	32	3.7	1243	2	US-08-897-340-5	Sequence 5, App1
741	32	3.7	970	3	US-09-248-335-51	Sequence 51, App1	814	32	3.7	1248	4	US-09-252-329-5	Sequence 5, App1
742	32	3.7	970	3	US-09-248-335-47	Sequence 47, App1	815	32	3.7	1248	4	US-09-020-956-171	Sequence 171, App1
743	32	3.7	972	1	US-07-915-934-1	Sequence 1, App1	816	32	3.7	1248	4	US-09-030-607-171	Sequence 171, App1
744	32	3.7	972	1	US-08-325-743-1	Sequence 1, App1	817	32	3.7	1248	4	US-09-605-785-171	Sequence 171, App1
745	32	3.7	972	4	US-09-549-831-5	Sequence 5, App1	818	32	3.7	1248	4	US-09-439-313-171	Sequence 171, App1
746	32	3.7	974	2	US-08-504-459-13	Sequence 13, App1	819	32	3.7	1248	4	US-09-352-616A-171	Sequence 171, App1
747	32	3.7	975	4	US-09-381-488-6	Sequence 6, App1	820	32	3.7	1248	4	US-09-232-149A-171	Sequence 171, App1
748	32	3.7	981	4	US-08-815-225-1	Sequence 1, App1	821	32	3.7	1275	3	US-08-725-532B-2	Sequence 2, App1
749	32	3.7	991	4	US-09-344-529-3	Sequence 3, App1	822	32	3.7	1279	3	US-09-248-335-25	Sequence 25, App1
750	32	3.7	996	2	US-08-624-650-2	Sequence 2, App1	823	32	3.7	1294	4	US-08-825-578-1	Sequence 1, App1
751	32	3.7	1001	1	US-08-728-259A-10	Sequence 10, App1	824	32	3.7	1296	4	US-08-425-578-1	Sequence 1, App1
752	32	3.7	1001	2	US-08-473-486-10	Sequence 10, App1	825	32	3.7	1296	4	US-08-772-402A-1	Sequence 1, App1
753	32	3.7	1003	2	US-08-887-997B-1	Sequence 1, App1	826	32	3.7	1298	1	US-08-361-708-1	Sequence 1, App1
754	32	3.7	1008	1	US-08-253-966B-15	Sequence 15, App1	827	32	3.7	1298	1	US-08-536-277-1	Sequence 1, App1
755	32	3.7	1009	4	US-09-291-992-15	Sequence 15, App1	828	32	3.7	1302	3	US-09-100-391-11	Sequence 11, App1
756	32	3.7	1013	1	US-07-920-519-30	Sequence 30, App1	829	32	3.7	1305	4	US-09-242-859A-3	Sequence 3, App1
757	32	3.7	1013	1	US-08-086-410-23	Sequence 23, App1	830	32	3.7	1305	4	US-09-242-859A-7	Sequence 7, App1

831	32	3.7	1307	2	US-08-960-022-17	Sequence 17, Appl	c 904	32	3.7	1472	4	US-09-006-595A-12	Sequence 12, Appl
832	32	3.7	1313	4	US-09-149-476-112	Sequence 112, App	905	32	3.7	1474	2	US-08-975-316-71	Sequence 71, Appl
833	32	3.7	1315	4	US-09-164-193-1	Sequence 1, Appl1	906	32	3.7	1474	4	US-09-615-192A-71	Sequence 71, Appl
834	32	3.7	1315	4	US-09-221-448A-1	Sequence 1, Appl1	907	32	3.7	1474	4	US-09-615-192A-402	Sequence 402, App
835	32	3.7	1319	2	US-08-504-459-7	Sequence 7, Appl1	908	32	3.7	1474	4	US-09-325-932A-202	Sequence 202, App
836	32	3.7	1325	1	US-08-306-691B-51	Sequence 51, Appl	909	32	3.7	1487	4	US-09-342-648-3	Sequence 3, Appl1
837	32	3.7	1325	2	US-08-464-517-1	Sequence 1, Appl1	910	32	3.7	1493	1	US-08-340-820-4	Sequence 24, Appl
838	32	3.7	1325	2	US-08-246-361A-1	Sequence 1, Appl1	911	32	3.7	1493	1	US-08-593-535-24	Sequence 24, Appl
839	32	3.7	1325	3	US-08-463-772-1	Sequence 1, Appl1	912	32	3.7	1501	6	5340934-5	Patent No. 5340934
840	32	3.7	1325	4	US-09-412-102-3	Sequence 3, Appl1	913	32	3.7	1503	3	US-08-923-454A-5	Sequence 5, Appl1
841	32	3.7	1325	4	US-09-217-787-3	Sequence 1, Appl1	914	32	3.7	1509	4	US-09-149-476-179	Sequence 179, App
842	32	3.7	1332	5	PCT-US93-05000-1	Sequence 3, Appl1	915	32	3.7	1511	2	US-08-809-763-3	Sequence 3, Appl1
843	32	3.7	1332	2	US-09-057-762-1	Sequence 1, Appl1	916	32	3.7	1511	2	US-08-956-253-3	Sequence 8, Appl1
844	32	3.7	1332	3	US-08-336-119A-1	Sequence 1, Appl1	917	32	3.7	1512	2	US-08-909-965C-8	Sequence 4, Appl1
845	32	3.7	1332	4	US-09-333-423-1	Sequence 1, Appl1	918	32	3.7	1512	4	US-09-157-603-4	Sequence 4, Appl1
846	32	3.7	1335	4	US-09-740-028A-1	Sequence 1, Appl1	919	32	3.7	1525	4	US-09-587-436-4	Sequence 4, Appl1
847	32	3.7	1338	2	US-08-484-938-7	Sequence 7, Appl1	920	32	3.7	1525	4	US-08-927-165A-4	Sequence 4, Appl1
848	32	3.7	1338	2	US-08-484-158B-7	Sequence 7, Appl1	921	32	3.7	1525	4	US-08-927-165A-6	Sequence 6, Appl1
849	32	3.7	1338	2	US-08-484-596A-7	Sequence 7, Appl1	922	32	3.7	1534	1	US-08-300-903A-6	Sequence 2, Appl1
850	32	3.7	1338	2	US-08-480-150A-7	Sequence 7, Appl1	923	32	3.7	1558	1	US-08-469-362-2	Sequence 2, Appl1
851	32	3.7	1338	3	US-08-458-731-7	Sequence 7, Appl1	924	32	3.7	1558	2	US-08-850-392-2	Sequence 2, Appl1
852	32	3.7	1338	3	US-08-149-223A-7	Sequence 7, Appl1	925	32	3.7	1559	2	US-08-417-174-1	Sequence 1, Appl1
853	32	3.7	1342	3	US-08-832-399-1	Sequence 1, Appl1	926	32	3.7	1559	2	US-08-231-565A-1	Sequence 1, Appl1
854	32	3.7	1342	4	US-09-372-498-1	Sequence 1, Appl1	927	32	3.7	1559	2	US-09-007-961-1	Sequence 1, Appl1
855	32	3.7	1342	4	US-09-286-529-18	Sequence 18, Appl	928	32	3.7	1559	2	US-08-267-439-1	Sequence 1, Appl1
856	32	3.7	1359	4	US-09-387-574-11	Sequence 11, Appl	929	32	3.7	1562	4	US-08-330-108-1	Sequence 1, Appl1
857	32	3.7	1359	4	US-09-668-096-11	Sequence 11, Appl	930	32	3.7	1562	5	PCT-US92-10087-1	Sequence 6, Appl1
858	32	3.7	1374	2	US-08-758-621-7	Sequence 7, Appl1	931	32	3.7	1576	1	US-08-157-101A-6	Sequence 1, Appl1
859	32	3.7	1374	4	US-09-107-858-7	Sequence 7, Appl1	932	32	3.7	1581	2	US-08-845-998-1	Sequence 1, Appl1
860	32	3.7	1375	4	US-09-372-422A-37	Sequence 37, Appl	933	32	3.7	1581	4	US-09-206-537-1	Sequence 1, Appl1
861	32	3.7	1378	4	US-09-149-476-208	Sequence 208, App	934	32	3.7	1581	4	US-09-313-300-1	Sequence 1, Appl1
862	32	3.7	1380	2	US-08-467-5598-1	Sequence 1, Appl1	935	32	3.7	1584	4	US-09-430-854-1	Sequence 1, Appl1
863	32	3.7	1384	4	US-09-227-357-142	Sequence 142, App	936	32	3.7	1584	4	US-08-745-995A-1	Sequence 3, Appl1
864	32	3.7	1386	2	US-08-897-340-1	Sequence 1, Appl1	937	32	3.7	1587	4	US-08-745-995A-3	Sequence 3, Appl1
865	32	3.7	1386	2	US-09-253-329-1	Sequence 1, Appl1	938	32	3.7	1587	4	US-09-724-864-5	Sequence 5, Appl1
866	32	3.7	1389	4	US-09-463-238-12	Sequence 12, Appl	939	32	3.7	1596	4	US-09-146-950-17	Sequence 17, Appl
867	32	3.7	1391	4	US-08-261-662-1	Sequence 1, Appl1	940	32	3.7	1600	4	US-09-091-097-33	Sequence 33, Appl
868	32	3.7	1391	5	PCT-US95-07752-1	Sequence 1, Appl1	941	32	3.7	1602	1	US-08-530-950-3	Sequence 3, Appl1
869	32	3.7	1393	1	US-08-174-467-18	Sequence 18, Appl	942	32	3.7	1602	3	US-08-888-429A-3	Sequence 3, Appl1
870	32	3.7	1393	3	US-08-452-071-18	Sequence 18, Appl	943	32	3.7	1602	4	US-09-149-879-3	Sequence 9, Appl1
871	32	3.7	1394	1	US-07-730-953-1	Sequence 1, Appl1	944	32	3.7	1605	4	US-09-091-097-9	Sequence 9, Appl1
872	32	3.7	1395	2	US-08-553-367A-1	Sequence 1, Appl1	945	32	3.7	1607	4	US-09-020-956-107	Sequence 107, App
873	32	3.7	1395	4	US-09-295-306-1	Sequence 1, Appl1	946	32	3.7	1621	4	US-09-030-607-107	Sequence 107, App
874	32	3.7	1395	4	US-09-734-719-1	Sequence 1, Appl1	947	32	3.7	1621	4	US-09-605-785-107	Sequence 107, App
875	32	3.7	1400	1	US-08-085-132-12	Sequence 12, Appl	948	32	3.7	1621	4	US-09-439-313-107	Sequence 107, App
876	32	3.7	1400	4	US-08-442-108B-22	Sequence 22, Appl	949	32	3.7	1621	4	US-09-352-616A-107	Sequence 107, App
877	32	3.7	1411	4	US-08-964-127-5	Sequence 5, Appl1	950	32	3.7	1621	4	US-09-334-601-9	Sequence 9, Appl1
878	32	3.7	1411	4	US-09-496-692-5	Sequence 5, Appl1	951	32	3.7	1621	4	US-08-583-318-4	Sequence 4, Appl1
879	32	3.7	1413	4	US-09-504-445-1	Sequence 1, Appl1	952	32	3.7	1622	1	US-08-737-524B-1	Sequence 1, Appl1
880	32	3.7	1420	4	US-09-227-357-21	Sequence 21, Appl	953	32	3.7	1631	2	US-08-665-037-1	Sequence 1, Appl1
881	32	3.7	1425	1	US-08-464-148-1	Sequence 1, Appl1	954	32	3.7	1631	2	US-08-665-067-1	Sequence 1, Appl1
882	32	3.7	1425	1	US-08-385-500-1	Sequence 1, Appl1	955	32	3.7	1639	2	US-08-732-870-1	Sequence 1, Appl1
883	32	3.7	1425	1	US-08-846-784-1	Sequence 1, Appl1	956	32	3.7	1642	2	US-09-443-041A-19	Sequence 19, Appl
884	32	3.7	1441	4	US-08-821-994-63	Sequence 63, Appl	957	32	3.7	1642	2	US-09-627-650B-2	Sequence 2, Appl1
885	32	3.7	1445	4	US-09-697-367-19	Sequence 19, Appl	958	32	3.7	1646	4	US-09-436-063C-2	Sequence 3, Appl1
886	32	3.7	1445	4	US-09-814-951A-1	Sequence 1, Appl1	959	32	3.7	1652	4	US-09-345-469-2	Sequence 2, Appl1
887	32	3.7	1447	4	US-09-443-041A-27	Sequence 27, Appl	960	32	3.7	1652	4	US-08-829-839-1	Sequence 1, Appl1
888	32	3.7	1448	4	US-09-387-212-12	Sequence 12, Appl	961	32	3.7	1654	4	US-08-991-426-3	Sequence 3, Appl1
889	32	3.7	1448	4	US-09-948-802-12	Sequence 12, Appl	962	32	3.7	1654	3	US-09-143-470-3	Sequence 3, Appl1
890	32	3.7	1454	2	US-08-713-000-7	Sequence 7, Appl1	963	32	3.7	1654	3	US-08-927-165A-4	Sequence 4, Appl1
891	32	3.7	1454	2	US-08-975-316-7	Sequence 7, Appl1	964	32	3.7	1664	4	US-08-709-838-1	Sequence 1, Appl1
892	32	3.7	1454	4	US-09-211-710-7	Sequence 7, Appl1	965	32	3.7	1670	3	US-08-829-839-1	Sequence 1, Appl1
893	32	3.7	1454	4	US-09-615-192A-7	Sequence 7, Appl1	966	32	3.7	1670	4	US-09-347-803-11	Sequence 11, Appl
894	32	3.7	1460	4	US-09-615-192A-95	Sequence 95, Appl	967	32	3.7	1683	1	US-08-143-219-26	Sequence 26, Appl
895	32	3.7	1466	4	US-08-984-919A-10	Sequence 10, Appl	968	32	3.7	1683	2	US-08-487-113D-118	Sequence 118, App
896	32	3.7	1466	4	US-08-984-919A-12	Sequence 12, Appl	969	32	3.7	1693	2	US-08-720-420A-118	Sequence 118, App
897	32	3.7	1469	3	US-09-074-912-1	Sequence 1, Appl1	970	32	3.7	1693	2	US-09-152-060-19	Sequence 19, Appl
898	32	3.7	1469	3	US-09-290-136-1	Sequence 1, Appl1	971	32	3.7	1699	4	US-09-000-062-1	Sequence 1, Appl1
899	32	3.7	1472	4	US-08-781-420-10	Sequence 10, Appl	972	32	3.7	1713	4	US-09-741-154-1	Sequence 139, Appl
900	32	3.7	1472	4	US-08-874-102-10	Sequence 10, Appl	973	32	3.7	1713	4	US-09-227-357-139	Sequence 139, Appl
901	32	3.7	1472	4	US-08-874-102-10	Sequence 10, Appl	974	32	3.7	1720	4	PCT-US96-00994-3	Sequence 4, Appl1
902	32	3.7	1472	4	US-09-006-595A-10	Sequence 10, Appl	975	32	3.7	1721	5		
903	32	3.7	1472	4	US-09-006-595A-10	Sequence 10, Appl	976	32	3.7	1721	5		

977	3.2	3.7	1.733	3	US-09-072-559-1	Sequence 1, Appl1
978	3.2	3.7	1.734	4	US-08-630-915A-23	Sequence 23, Appl1
979	3.2	3.7	1.741	1	US-08-566-655-5	Sequence 5, Appl1
980	3.2	3.7	1.745	4	US-09-388-349-5	Sequence 5, Appl1
981	3.2	3.7	1.747	1	US-08-176-620A-1	Sequence 1, Appl1
982	3.2	3.7	1.747	1	US-08-463-862-1	Sequence 1, Appl1
983	3.2	3.7	1.747	2	US-08-461-985-1	Sequence 1, Appl1
984	3.2	3.7	1.747	2	US-08-456-887-1	Sequence 1, Appl1
985	3.2	3.7	1.747	2	US-08-932-787B-1	Sequence 1, Appl1
986	3.2	3.7	1.747	4	US-08-932-012C-1	Sequence 1, Appl1
987	3.2	3.7	1.747	4	US-08-888-818C-1	Sequence 1, Appl1
988	3.2	3.7	1.772	2	US-08-960-022-13	Sequence 13, Appl1
989	3.2	3.7	1.780	4	US-09-204-548B-5	Sequence 5, Appl1
990	3.2	3.7	1.781	1	US-08-314-615-2	Sequence 2, Appl1
991	3.2	3.7	1.781	1	US-08-433-362-2	Sequence 2, Appl1
992	3.2	3.7	1.781	1	US-08-245-295-4	Sequence 4, Appl1
993	3.2	3.7	1.781	1	US-08-245-295-4	Sequence 4, Appl1
994	3.2	3.7	1.781	1	US-08-466-984A-4	Sequence 4, Appl1
995	3.2	3.7	1.781	1	US-08-656-604-4	Sequence 4, Appl1
996	3.2	3.7	1.781	1	US-08-482-882-2	Sequence 2, Appl1
997	3.2	3.7	1.781	1	US-08-485-604-4	Sequence 2, Appl1
998	3.2	3.7	1.781	1	US-08-483-389-2	Sequence 2, Appl1
999	3.2	3.7	1.781	1	US-08-487-113D-2	Sequence 2, Appl1
1000	3.2	3.7	1.781	2	US-08-487-595-4	Sequence 4, Appl1

## ALIGNMENTS

RESULT 1  
US-08-706-214-2  
Sequence 2, Application US/08706214  
Patent No. 5770720  
GENERAL INFORMATION:  
APPLICANT: Deul, Thomas F.  
APPLICANT: Wang, Zhao-Yi  
APPLICANT: Shenk, Thomas E.  
TITLE OF INVENTION: UNBQUITTIN CONJUGATING ENZYMES HAVING  
TITLE OF INVENTION: TRANSCRIPTIONAL REPRESSOR ACTIVITY  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Paul A. Stone, Esq.  
STREET: One Metropolitan Square, 16th Floor  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/706,214  
FILING DATE: 30-AUG-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/018,040  
FILING DATE: 21-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/002,995  
FILING DATE: 30-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Stone, Paul A.  
REGISTRATION NUMBER: 38,628  
REFERENCE/DOCKET NUMBER: JWH 10033  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 231-4340  
TELEFAX: 6502697583 MCI  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1137 base pairs

```

; type: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
HYPOTHEITICAL: NO
; ANTI-SENSE: NO
FEATURE:
; NAME/KEY: CDS
; LOCATION: 88..564
US-08-706-214-2

Query Match              4.3%; Score 37; DB 1; Length 1137;
Best Local Similarity    100.0%; Pred.No. 2.5e-07;
Matches   37; Conservative      0; Mismatches     0; Indels      0; Gaps       0;

Oy      827 TATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 863
||| ||||||| |
Db      1100 TATTATAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1136
||||| |

RESULT 2
US-08-706-214-1
; Sequence 1, Application US/08706214
; Patent No. 5770720
```

```

RESULT 2
US-08-706-214-1
Sequence 1, Application US/08706214
Patent No. 570720
GENERAL INFORMATION:
APPLICANT: Deul, Thomas F.
APPLICANT: Wang, Zhao-Yi
APPLICANT: Shenk, Thomas E.
TITLE OF INVENTION: UBIQUITIN CONJUGATING ENZYMES HAVING
TITLE OF INVENTION: TRANSCRIPTIONAL REPRESSOR ACTIVITY
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Paul A. Stone, Esq.
STREET: One Metropolitan Square, 16th Floor
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/706,214
FILING DATE: 30-AUG-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/018,040
FILING DATE: 21-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/002,995
FILING DATE: 30-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Stone, Paul A.
REGISTRATION NUMBER: 38,628
REFERENCE/DOCKET NUMBER: JWH 10033
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 231-5400
TELEFAX: (314) 231-4342
TELEX: 6502697583 MCI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1856 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURES:
NAME/KEY: CDS
LOCATION: 807..1283

```











Mon Feb 10 10:37:37 2003

us-10-021-811-35.oligo.m1

Job time : 90 secs

0

2

1



80	39	4.5	1101	17	CNS00CPQ	AL059482 Drosophila	153	37	4.3	140	12	BG454124	BG454124 NP096G06L
C 81	39	4.5	1105	12	BE965491	BE965491 601659420	C 154	37	4.3	144	9	AT262151	AT262151 qz28e03.x
C 82	39	4.5	1405	13	BI489285	BI489285 603021628	C 155	37	4.3	144	9	AT734004	AT734004 qz28e03.x
C 83	39	4.5	1418	12	BG030516	BG030516 602288054	C 156	37	4.3	146	9	AT734004	AT734004 qz28e03.x
C 84	38	4.4	62	9	AA398875	AA398875 v071d04.x	C 157	37	4.3	148	12	BF750886	BF750886 RCI-BN042
C 85	38	4.4	100	13	BI001553	BI001553 MR3-HN014	C 158	37	4.3	152	14	BQ393689	BQ393689 RCI-BN042
C 86	38	4.4	134	17	CNS01553	AL241313 Tetrachodon	C 159	37	4.3	156	14	C25761	C25761 C25761
C 87	38	4.4	140	13	BG977694	BG977694 MR2-CI018	C 160	37	4.3	166	9	AI915279	AI915279 tx06e01.x
C 88	38	4.4	150	10	BE069307	BE069307 QV3-BT038	C 161	37	4.3	169	9	AI249689	AI249689 tx06e01.x
C 89	38	4.4	148	10	BE248148	BE248148 NF002A07D	C 162	37	4.3	174	10	BB588786	BB588786 tx06e01.x
C 90	38	4.4	180	13	BQ326229	BQ326229 BU332629	C 163	37	4.3	176	9	AI699784	AI699784 tx21e04.x
C 91	38	4.4	236	10	AM432525	AM432525 sb75d08.y	C 164	37	4.3	189	13	BM270126	BM270126 sak24910.
C 92	38	4.4	240	9	AU074936	AU074936 AU074936	C 165	37	4.3	191	12	BE804703	BE804703 sr78e06.y
C 93	38	4.4	242	13	BI502406	BI502406 CT49C10.y	C 166	37	4.3	192	10	AM306966	AM306966 sf51a10.y
C 94	38	4.4	248	13	BI704970	BI704970 tr60909.y	C 167	37	4.3	194	12	BG237216	BG237216 saag8c08.
C 95	38	4.4	251	12	BF807864	BF807864 RC3-CI004	C 168	37	4.3	197	9	AI631876	AI631876 wa37b12.x
C 96	38	4.4	272	9	AL677222	AL677222	C 169	37	4.3	207	9	AI323848	AI323848 mt66g09.x
C 97	38	4.4	275	13	BM346103	BM346103 tr12d02.y	C 170	37	4.3	208	13	BM092623	BM092623 sbh15606.
C 98	38	4.4	275	14	BQ455797	BQ455797 ke22h12.y	C 171	37	4.3	209	13	BM092623	BM092623 sbh15606.
C 99	38	4.4	278	14	BQ612645	BQ612645 sap72h02.	C 172	37	4.3	215	13	BG991154	BG991154
C 100	38	4.4	300	9	AU263329	AU263329	C 173	37	4.3	218	10	BE624780	BE624780
C 101	38	4.4	303	13	BG939743	BG939743 cr61e04.x	C 174	37	4.3	219	9	AI919434	AI919434
C 102	38	4.4	317	13	BM280884	BM280884 KI06d06.y	C 175	37	4.3	221	12	BG001296	BG001296
C 103	38	4.4	326	10	AM785659	AM785659 117007.MA	C 176	37	4.3	222	14	C92990	C92990
C 104	38	4.4	333	9	AU261286	AU261286	C 177	37	4.3	228	12	BG651671	BG651671
C 105	38	4.4	343	9	BQ748672	BQ748672	C 178	37	4.3	237	9	AA576046	AA576046
C 106	38	4.4	349	14	BM021890	BM021890	C 179	37	4.3	240	9	AI824522	AI824522
C 107	38	4.4	352	9	BG450094	BG450094 NF011C09D	C 180	37	4.3	240	12	BG841917	BG841917
C 108	38	4.4	358	12	BG450094	BG450094 NF011C09D	C 181	37	4.3	244	10	AM666166	AM666166
C 109	38	4.4	370	9	AL803313	AL803313	C 182	37	4.3	254	9	AU094219	AU094219
C 110	38	4.4	378	10	AM990222	AM990222	C 183	37	4.3	254	9	AU094219	AU094219
C 111	38	4.4	380	11	BM027649	BM027649 Mus muscu	C 184	37	4.3	259	13	BI001794	BI001794
C 112	38	4.4	386	14	C94469	C94469 C94469 Dict	C 185	37	4.3	260	14	BQ423248	BQ423248
C 113	38	4.4	399	13	BG965590	BG965590 602830619	C 186	37	4.3	260	14	BQ423248	BQ423248
C 114	38	4.4	429	14	BQ748073	BQ748073 UI-M-FA0-	C 187	37	4.3	270	9	AL635155	AL635155
C 115	38	4.4	441	10	AM635529	AM635529 b140c02.w	C 188	37	4.3	274	12	BF634988	BF634988
C 116	38	4.4	445	10	AM635581	AM635581 b140c02.w	C 189	37	4.3	276	10	AM423760	AM423760
C 117	38	4.4	455	13	BG926629	BG926629 HNC33-1-D	C 190	37	4.3	281	13	BM521098	BM521098
C 118	38	4.4	466	9	AL803394	AL803394	C 191	37	4.3	282	13	BI454782	BI454782
C 119	38	4.4	471	13	BI945249	BI945249 sb4se12.y	C 192	37	4.3	288	13	BI454782	BI454782
C 120	38	4.4	500	10	BE693203	BE693203 zewp0107c	C 193	37	4.3	292	10	AM099060	AM099060
C 121	38	4.4	516	12	BE893404	BE893404 601436735	C 194	37	4.3	292	10	AM099060	AM099060
C 122	38	4.4	524	12	BG202913	BG202913 RST2281	C 195	37	4.3	294	13	BI742580	BI742580
C 123	38	4.4	555	10	BE408765	BE408765 601303457	C 196	37	4.3	294	13	BI742580	BI742580
C 124	38	4.4	576	12	BE965901	BE965901 601659080	C 197	37	4.3	296	9	AI418984	AI418984
C 125	38	4.4	578	14	BQ523351	BQ523351 NISC. n121	C 198	37	4.3	297	9	AI860619	AI860619
C 126	38	4.4	616	2	BM830863	BM830863 K-EST0104	C 199	37	4.3	298	10	AM56318	AM56318
C 127	38	4.4	628	9	AL667177	AL667177	C 200	37	4.3	302	12	BG237811	BG237811
C 128	38	4.4	638	17	AQ329262	AQ329262 NDXB0044N	C 201	37	4.3	308	14	BQ621347	BQ621347
C 129	38	4.4	675	10	BE573505	BE573505 601332513	C 202	37	4.3	315	12	BF584526	BF584526
C 130	38	4.4	755	13	BI489221	BI489221 603021446	C 203	37	4.3	317	14	C90426	C90426
C 131	38	4.4	758	10	BE254113	BE254113 601113971	C 204	37	4.3	322	9	AI253089	AI253089
C 132	38	4.4	764	10	AM332484	AM332484 S9A9 AGS-	C 205	37	4.3	322	9	AI253089	AI253089
C 133	38	4.4	766	11	BC018268	BC018268 Mus muscu	C 206	37	4.3	329	9	AI253089	AI253089
C 134	38	4.4	832	12	BE729660	BE729660 601562804	C 207	37	4.3	332	9	AI253089	AI253089
C 135	38	4.4	853	17	AZ545711	AZ545711 ENTGDI9TF	C 208	37	4.3	332	9	AI253089	AI253089
C 136	38	4.4	887	12	BF274559	BF274559 GA-ED002	C 209	37	4.3	336	14	BM924352	BM924352
C 137	38	4.4	925	13	BM416279	BM416279 OP21369 M	C 210	37	4.3	336	14	BM924352	BM924352
C 138	38	4.4	960	17	AG031602	AG031602 Pan trogl	C 211	37	4.3	336	14	BM924352	BM924352
C 139	38	4.4	985	13	BM415996	BM415996 OP21079 M	C 212	37	4.3	340	10	AM172850	AM172850
C 140	38	4.4	1000	13	BI521761	BI521761 603081730	C 213	37	4.3	345	10	AM172850	AM172850
C 141	38	4.4	1012	14	BQ433017	BQ433017 AGENCOURT	C 214	37	4.3	348	10	AM681818	AM681818
C 142	38	4.4	1090	14	BQ717503	BQ717503 AGENCOURT	C 215	37	4.3	354	9	AU057808	AU057808
C 143	38	4.4	1101	17	CNS01406	AL104616 Drosophila	C 216	37	4.3	360	9	AU183904	AU183904
C 144	38	4.4	1161	12	BQ479771	BQ479771 602526978	C 217	37	4.3	363	14	C24403	C24403
C 145	38	4.4	1363	12	BE965813	BE965813 601655584	C 218	37	4.3	376	10	AM496734	AM496734
C 146	38	4.4	1750	13	BQ911020	BQ911020 602806383	C 219	37	4.3	385	14	BQ393689	BQ393689
C 147	37	4.3	72	14	T25592	T25592 EST00628 Un	C 220	37	4.3	386	13	BI524350	BI524350
C 148	37	4.3	93	13	BQ963829	BQ963829 602831491	C 221	37	4.3	388	10	AM506617	AM506617
C 149	37	4.3	111	14	BM966736	BM966736 ko11b07.y	C 222	37	4.3	398	10	AI816490	AI816490
C 150	37	4.3	124	9	AA616273	AA616273 vnf64a07.x	C 223	37	4.3	401	9	AU098348	AU098348
C 151	37	4.3	126	13	BM541971	BM541971 AGENCOURT	C 224	37	4.3	402	12	BG511676	BG511676
C 152	37	4.3	137	12	BF644938	BF644938 NF028A09E	C 225	37	4.3				

C 226	37	4.3	411	9	AI537079	AI537079 t082h04.x	C 229	36	4.2	106	9	AA490906	AA490906 aa52b09.s
227	37	4.3	411	13	BM521226	BM521226 sa112b08.	C 300	36	4.2	111	14	BQ225032	BQ225032 ACENOCURT
C 228	37	4.3	415	10	AV657079	AV657079	C 301	36	4.2	112	13	BG956245	BG956245 OVI-CT075
C 229	37	4.3	416	9	AL803452	AL803452	C 302	36	4.2	120	13	BM478091	BM478091 ACENOCURT
C 230	37	4.3	424	9	AU262968	AU262968	C 303	36	4.2	121	12	BF915483	BF915483 MR3-UT012
231	37	4.3	428	12	BE881033	BE881033 601490790	C 304	36	4.2	123	13	BT742035	BT742035 Kc89p07.y
C 232	37	4.3	440	13	BI936227	BI936227 PFEST0a2	C 305	36	4.2	125	9	AU071541	AU071541 AU071541
C 233	37	4.3	464	9	AA659923	AA659923 nv03e08.s	C 306	36	4.2	128	9	AU073398	AU073398 AU073398
C 234	37	4.3	464	9	AI138585	AI138585 qq93f08.x	C 307	36	4.2	131	14	BQ237436	BQ237436 TAE05016G
C 235	37	4.3	464	9	AU262244	AU262244 AU262244	C 308	36	4.2	133	13	BI536278	BI536278 390802 MA
236	37	4.3	472	14	C93369	C93369 C93369 D1ct	C 309	36	4.2	135	12	BF338132	BF338132 602037926
C 237	37	4.3	477	9	AL636291	AL636291	C 310	36	4.2	137	13	BG950524	BG950524 PM0-CT070
C 238	37	4.3	484	10	AM584191	AM584191 N210131e	C 311	36	4.2	141	13	BJ427681	BJ427681 BJ427681
C 239	37	4.3	486	10	AM544452	AM544452 C0181h12-	C 312	36	4.2	142	13	BJ337720	BJ337720 BJ337720
C 240	37	4.3	492	10	AM641185	AM641185 cm04C09.w	C 313	36	4.2	143	13	BJ331067	BJ331067 BJ331067
C 241	37	4.3	494	9	AL641624	AL641624 AL641624	C 314	36	4.2	144	9	AI944844	AI944844 bs06b03.y
C 242	37	4.3	494	9	BO395433	BO395433 NISC_nq15	C 315	36	4.2	144	13	BJ395033	BJ395033 BJ395033
C 243	37	4.3	518	14	BO527377	BO527377 NISC_nq22	C 316	36	4.2	144	13	BJ397854	BJ397854 BJ397854
C 244	37	4.3	524	12	BF015608	BF015608 uy23g02.y	C 317	36	4.2	145	9	AI440260	AI440260 tJ01d05.x
C 245	37	4.3	530	9	AI581183	AI581183 t194f03.x	C 318	36	4.2	145	10	AM169379	AM169379 KJ25f06.x
C 246	37	4.3	533	9	AI691234	AI691234 606025H09	C 319	36	4.2	146	13	BI053612	BI053612 MR3-GN046
247	37	4.3	547	9	AU261551	AU261551 AU261551	C 320	36	4.2	147	9	AI340603	AI340603 cb31g08.x
C 248	37	4.3	552	9	AI195074	AI195074 MA002345.	C 321	36	4.2	147	9	AI960330	AI960330 sc82a06.y
C 249	37	4.3	566	14	BO526714	BO526714 NISC_no18	C 322	36	4.2	149	12	BE976721	BE976721 bs55a08.y
C 250	37	4.3	568	14	BO396965	BO396965 NISC_nq25	C 323	36	4.2	150	9	AI267162	AI267162 aq49f03.x
C 251	37	4.3	570	14	BO522054	BO522054 NISC_n114	C 324	36	4.2	150	9	AI965757	AI965757 sc75f09.y
C 252	37	4.3	577	14	BO395988	BO395988 NISC_nq18	C 325	36	4.2	150	13	BJ420264	BJ420264 BJ420264
C 253	37	4.3	580	14	BO393588	BO393588 NISC_nq04	C 326	36	4.2	151	9	AI432935	AI432935 tH44d08.x
C 254	37	4.3	582	14	BO523851	BO523851 NISC_no01	C 327	36	4.2	151	10	AM008265	AM008265 wv52e07.x
C 255	37	4.3	590	14	BO524387	BO524387 NISC_no04	C 328	36	4.2	151	13	BJ337913	BJ337913 BJ337913
C 256	37	4.3	592	10	BE038678	BE038678 AB04A02 A	C 329	36	4.2	152	13	BI521790	BI521790 603081775
C 257	37	4.3	600	14	BO396574	BO396574 NISC_nq21	C 330	36	4.2	152	13	BJ364859	BJ364859 BJ364859
C 258	37	4.3	607	14	BO415149	BO415149 GA_Ed009	C 331	36	4.2	152	13	BJ366714	BJ366714 BJ366714
C 259	37	4.3	622	10	BM0130525	BM0130525 xff52e06.x	C 332	36	4.2	152	13	BJ397535	BJ397535 BJ397535
C 260	37	4.3	623	13	BM013585	BM013585 603639061	C 333	36	4.2	153	9	AI037454	AI037454 DKF2P762K
C 261	37	4.3	629	14	BO389845	BO389845 NISC_mq09	C 334	36	4.2	153	9	AI138406	AI138406 DXF2P762K
C 262	37	4.3	629	14	BO527405	BO527405 NISC_no22	C 335	36	4.2	153	13	BI943053	BI943053 sn09D06.y
263	37	4.3	634	9	AU037693	AU037693 AU037693	C 336	36	4.2	153	13	BJ386665	BJ386665 BJ386665
264	37	4.3	638	13	BI524025	BI524025 603052190	C 337	36	4.2	153	13	BJ393754	BJ393754 BJ393754
265	37	4.3	642	13	BI822759	BI822759 603040213	C 338	36	4.2	153	13	BJ418243	BJ418243 BJ418243
C 266	37	4.3	644	10	AM61968	AM61968 T8236 MVA	C 339	36	4.2	154	13	BI783472	BI783472 kJ20g05.y
C 267	37	4.3	668	10	AM766225	AM766225 da83h05.y	C 340	36	4.2	154	13	BJ422991	BJ422991 BJ422991
C 268	37	4.3	669	10	AV706671	AV706671	C 341	36	4.2	155	12	BE964614	BE964614 601658540
C 269	37	4.3	677	13	BI490413	BI490413 603031660	C 342	36	4.2	155	13	BJ330298	BJ330298 BJ330298
C 270	37	4.3	717	10	AV716471	AV716471	C 343	36	4.2	155	13	BJ333306	BJ333306 BJ333306
C 271	37	4.3	720	12	BG289378	BG289378 602381405	C 344	36	4.2	155	13	BJ365333	BJ365333 BJ365333
C 272	37	4.3	732	12	BG114012	BG114012 602284496	C 345	36	4.2	155	13	BJ393309	BJ393309 BJ393309
C 273	37	4.3	738	12	BF300482	BF300482 602031614	C 346	36	4.2	155	13	BJ426387	BJ426387 BJ426387
C 274	37	4.3	742	13	BI490154	BI490154 603031905	C 347	36	4.2	155	17	N97520	N97520 0620M3 gmbP
275	37	4.3	744	11	AV106416	AV106416 Zea maye	C 348	36	4.2	156	13	BJ333618	BJ333618 BJ333618
276	37	4.3	774	10	AV755874	AV755874	C 349	36	4.2	156	13	BJ369364	BJ369364 BJ369364
277	37	4.3	805	13	BI091593	BI091593 6029859015	C 350	36	4.2	156	13	BJ425228	BJ425228 BJ425228
278	37	4.3	822	12	BF970631	BF970631 602273968	C 351	36	4.2	157	13	BI022366	BI022366 CM3-WT034
279	37	4.3	826	13	BI812817	BI812817 603082632	C 352	36	4.2	157	13	BJ334481	BJ334481 BJ334481
C 280	37	4.3	937	17	CNS006ST	CNS006ST	C 353	36	4.2	157	13	BJ335390	BJ335390 BJ335390
C 281	37	4.3	1034	12	BE962970	BE962970 601656332	C 354	36	4.2	157	13	BJ367353	BJ367353 BJ367353
C 282	37	4.3	1213	10	BE407722	BE407722 HMM001.HO	C 355	36	4.2	157	13	BJ368954	BJ368954 BJ368954
C 283	37	4.3	1250	13	BI489117	BI489117 603021186	C 356	36	4.2	157	13	BJ393595	BJ393595 BJ393595
C 284	37	4.3	1258	12	BE955881	BE955881 601659056	C 357	36	4.2	157	13	BJ394883	BJ394883 BJ394883
C 285	37	4.3	1284	13	BI519336	BI519336 603062208	C 358	36	4.2	158	13	BJ335804	BJ335804 BJ335804
C 286	37	4.3	1304	12	BE967292	BE967292 601661162	C 359	36	4.2	158	13	BJ365086	BJ365086 BJ365086
C 287	37	4.3	1324	12	BE732599	BE732599 601571156	C 360	36	4.2	158	13	BJ366675	BJ366675 BJ366675
C 288	37	4.3	1387	12	BF136629	BF136629 601903179	C 361	36	4.2	158	13	BJ367046	BJ367046 BJ367046
C 289	37	4.3	1474	12	BG261116	BG261116 602372746	C 362	36	4.2	158	13	BJ368799	BJ368799 BJ368799
C 290	37	4.3	1589	12	BG419146	BG419146 602446384	C 363	36	4.2	159	13	BJ396125	BJ396125 BJ396125
C 291	37	4.3	1625	12	BE963209	BE963209 601656826	C 364	36	4.2	159	13	BJ330804	BJ330804 BJ330804
C 292	37	4.3	2025	12	BE963127	BE963127 601656929	C 365	36	4.2	159	13	BJ331326	BJ331326 BJ331326
C 293	37	4.3	2182	12	BF981953	BF981953 602306101	C 366	36	4.2	159	13	BJ336432	BJ336432 BJ336432
C 294	36	4.2	60	9	AA422440	AA422440 vrl12d02.s	C 367	36	4.2	159	13	BJ365763	BJ365763 BJ365763
C 295	36	4.2	87	3	BI495570	BI495570 dfl19b09.	C 368	36	4.2	159	13	BJ394806	BJ394806 BJ394806
C 296	36	4.2	88	9	AU261695	AU261695 AU261695	C 369	36	4.2	159	13	BJ424253	BJ424253 BJ424253
C 297	36	4.2	100	12	BF815228	BF815228 MR2-CT012	C 370	36	4.2	160	13	BE965724	BE965724 601659783
C 298	36	4.2	105	10	AW091417	AW091417 683023E11	C 371	36	4.2	160	13	BJ332589	BJ332589 BJ332589

372	C	372	4.2	160	13	BJ334258	BJ334258	BJ334258	C	445	4.2	228	14	BM527688	BM527688	NTSC no23
373	C	373	4.2	160	13	BJ395162	BJ395162	BJ395162	C	445	4.2	229	10	AM600647	AM600647	NTSC no23
374	C	374	4.2	161	13	BJ368633	BJ368633	BJ368633	C	447	4.2	229	12	BF192242	BF192242	NTSC no23
375	C	375	4.2	166	14	BQ389420	BQ389420	BQ389420	C	448	4.2	230	13	BM030846	BM030846	NTSC no23
376	C	376	4.2	169	13	BQ38387	BQ38387	BQ38387	C	449	4.2	231	9	AU263276	AU263276	NTSC no23
377	C	377	4.2	170	13	BQ38387	BQ38387	BQ38387	C	449	4.2	231	13	BM1746763	BM1746763	NTSC no23
378	C	378	4.2	170	13	BQ38387	BQ38387	BQ38387	C	450	4.2	231	13	BM1746763	BM1746763	NTSC no23
379	C	379	4.2	171	13	BQ38387	BQ38387	BQ38387	C	451	4.2	231	17	CNS03D27	CNS03D27	NTSC no23
380	C	380	4.2	171	13	BQ38387	BQ38387	BQ38387	C	452	4.2	232	13	N98035	N98035	NTSC no23
381	C	381	4.2	172	13	BM092383	BM092383	BM092383	C	453	4.2	233	13	BQ339101	BQ339101	NTSC no23
382	C	382	4.2	173	9	AL640567	AL640567	AL640567	C	454	4.2	234	14	BM874445	BM874445	NTSC no23
383	C	383	4.2	174	10	AV962602	AV962602	AV962602	C	455	4.2	235	10	AM129604	AM129604	NTSC no23
384	C	384	4.2	174	10	AM187209	AM187209	AM187209	C	456	4.2	236	13	BM1744868	BM1744868	NTSC no23
385	C	385	4.2	177	13	BQ394702	BQ394702	BQ394702	C	457	4.2	240	13	BM506431	BM506431	NTSC no23
386	C	386	4.2	178	9	AL588434	AL588434	AL588434	C	458	4.2	241	9	AL686596	AL686596	NTSC no23
387	C	387	4.2	179	12	BQ237446	BQ237446	BQ237446	C	459	4.2	245	12	BM003106	BM003106	NTSC no23
388	C	388	4.2	180	13	BI894908	BI894908	BI894908	C	460	4.2	245	13	BM154504	BM154504	NTSC no23
389	C	389	4.2	180	13	BM091809	BM091809	BM091809	C	461	4.2	247	9	AT080087	AT080087	NTSC no23
390	C	390	4.2	181	13	BQ933681	BQ933681	BQ933681	C	462	4.2	248	14	BM889429	BM889429	NTSC no23
391	C	391	4.2	183	12	BF923395	BF923395	BF923395	C	463	4.2	249	12	BF722381	BF722381	NTSC no23
392	C	392	4.2	184	9	AL666430	AL666430	AL666430	C	464	4.2	250	10	AM187451	AM187451	NTSC no23
393	C	393	4.2	184	10	AM458261	AM458261	AM458261	C	465	4.2	250	12	BQ736841	BQ736841	NTSC no23
394	C	394	4.2	184	10	BE904911	BE904911	BE904911	C	466	4.2	250	14	BM900253	BM900253	NTSC no23
395	C	395	4.2	184	14	C84800	C84800	C84800	C	467	4.2	251	9	AA911635	AA911635	NTSC no23
396	C	396	4.2	185	9	AL567883	AL567883	AL567883	C	468	4.2	257	12	BM237821	BM237821	NTSC no23
397	C	397	4.2	185	13	BI042555	BI042555	BI042555	C	470	4.2	257	14	BQ666875	BQ666875	NTSC no23
398	C	398	4.2	186	14	BQ567326	BQ567326	BQ567326	C	471	4.2	258	14	BM881955	BM881955	NTSC no23
400	C	400	4.2	188	9	AL640991	AL640991	AL640991	C	472	4.2	259	14	N97931	N97931	NTSC no23
401	C	401	4.2	188	13	BQ30659	BQ30659	BQ30659	C	473	4.2	261	14	BM896676	BM896676	NTSC no23
402	C	402	4.2	189	13	BI15626	BI15626	BI15626	C	474	4.2	262	13	BI745508	BI745508	NTSC no23
403	C	403	4.2	190	13	BM143922	BM143922	BM143922	C	475	4.2	265	9	AU262271	AU262271	NTSC no23
404	C	404	4.2	191	14	BQ739553	BQ739553	BQ739553	C	476	4.2	265	13	BM186101	BM186101	NTSC no23
405	C	405	4.2	193	9	AA586749	AA586749	AA586749	C	477	4.2	266	13	BI863396	BI863396	NTSC no23
406	C	406	4.2	193	12	BF916550	BF916550	BF916550	C	478	4.2	267	10	AM301863	AM301863	NTSC no23
407	C	407	4.2	194	17	CNS072MX	CNS072MX	CNS072MX	C	479	4.2	267	12	BQ736549	BQ736549	NTSC no23
408	C	408	4.2	196	9	AL836066	AL836066	AL836066	C	480	4.2	268	14	BM774871	BM774871	NTSC no23
409	C	409	4.2	199	13	BQ333894	BQ333894	BQ333894	C	481	4.2	269	12	BF753033	BF753033	NTSC no23
410	C	410	4.2	200	12	BG606920	BG606920	BG606920	C	482	4.2	270	9	AL661209	AL661209	NTSC no23
411	C	411	4.2	201	9	AA151109	AA151109	AA151109	C	483	4.2	274	10	BE233379	BE233379	NTSC no23
412	C	412	4.2	202	9	AU038819	AU038819	AU038819	C	484	4.2	274	13	BM181089	BM181089	NTSC no23
413	C	413	4.2	202	10	BE255229	BE255229	BE255229	C	485	4.2	276	14	C84796	C84796	NTSC no23
414	C	414	4.2	202	13	BQ333902	BQ333902	BQ333902	C	486	4.2	276	13	BQ395509	BQ395509	NTSC no23
415	C	415	4.2	203	13	BQ330915	BQ330915	BQ330915	C	487	4.2	278	14	BM880671	BM880671	NTSC no23
416	C	416	4.2	203	13	BQ330921	BQ330921	BQ330921	C	488	4.2	279	10	BE033059	BE033059	NTSC no23
417	C	417	4.2	203	13	BQ333375	BQ333375	BQ333375	C	489	4.2	280	13	BM531944	BM531944	NTSC no23
418	C	418	4.2	203	13	BQ333375	BQ333375	BQ333375	C	490	4.2	282	12	BQ736615	BQ736615	NTSC no23
419	C	419	4.2	204	12	BQ237154	BQ237154	BQ237154	C	491	4.2	282	13	BI902697	BI902697	NTSC no23
420	C	420	4.2	204	13	BQ333273	BQ333273	BQ333273	C	492	4.2	285	12	BF059761	BF059761	NTSC no23
421	C	421	4.2	205	14	BQ629820	BQ629820	BQ629820	C	493	4.2	285	13	BI745476	BI745476	NTSC no23
422	C	422	4.2	206	9	AL718325	AL718325	AL718325	C	494	4.2	287	12	BF965814	BF965814	NTSC no23
423	C	423	4.2	207	13	BM635418	BM635418	BM635418	C	495	4.2	288	13	BI268512	BI268512	NTSC no23
424	C	424	4.2	208	10	BE056103	BE056103	BE056103	C	496	4.2	288	13	BI863466	BI863466	NTSC no23
425	C	425	4.2	209	13	BM087314	BM087314	BM087314	C	497	4.2	288	14	BQ527692	BQ527692	NTSC no23
426	C	426	4.2	210	9	AL640398	AL640398	AL640398	C	498	4.2	289	9	AL660133	AL660133	NTSC no23
427	C	427	4.2	210	12	BF069697	BF069697	BF069697	C	499	4.2	289	10	AM429843	AM429843	NTSC no23
428	C	428	4.2	210	13	BI321301	BI321301	BI321301	C	500	4.2	290	13	BI745120	BI745120	NTSC no23
429	C	429	4.2	211	13	BI945109	BI945109	BI945109	C	501	4.2	291	14	BQ640732	BQ640732	NTSC no23
430	C	430	4.2	214	10	AM193975	AM193975	AM193975	C	502	4.2	291	13	BI408847	BI408847	NTSC no23
431	C	431	4.2	215	10	AMW87457	AMW87457	AMW87457	C	503	4.2	294	13	BI747245	BI747245	NTSC no23
432	C	432	4.2	217	10	BE059127	BE059127	BE059127	C	504	4.2	295	13	BM345035	BM345035	NTSC no23
433	C	433	4.2	217	13	BI865874	BI865874	BI865874	C	505	4.2	297	13	BQ662133	BQ662133	NTSC no23
434	C	434	4.2	219	13	BI541562	BI541562	BI541562	C	506	4.2	298	13	BQ941170	BQ941170	NTSC no23
435	C	435	4.2	219	13	BI781603	BI781603	BI781603	C	507	4.2	299	9	AL595093	AL595093	NTSC no23
436	C	436	4.2	221	9	AL165384	AL165384	AL165384	C	508	4.2	299	13	BI501913	BI501913	NTSC no23
437	C	437	4.2	221	13	BQ953677	BQ953677	BQ953677	C	509	4.2	300	9	AL047039	AL047039	NTSC no23
438	C	438	4.2	221	13	BI042928	BI042928	BI042928	C	510	4.2	302	9	AL514717	AL514717	NTSC no23
439	C	439	4.2	222	10	BM102605	BM102605	BM102605	C	511	4.2	302	9	AL1918484	AL1918484	NTSC no23
440	C	440	4.2	222	13	AM101889	AM101889	AM101889	C	512	4.2	312	12	BQ457790	BQ457790	NTSC no23
441	C	441	4.2	223	13	BI744972	BI744972	BI744972	C	513	4.2	313	9	AL012928	AL012928	NTSC no23
442	C	442	4.2	223	13	BM124451	BM124451	BM124451	C	514	4.2	314	13	BM311810	BM311810	NTSC no23
443	C	443	4.2	226	14	BM966662	BM966662	BM966662	C	515	4.2	315	12	BG178237	BG178237	NTSC no23
444	C	444	4.2	228	13	BG993456	BG993456	BG993456	C	517	4.2	315	12	BG736327	BG736327	NTSC no23

518	36	4.2	316	10	AM597643	AM597643 sj49a06.y	591	36	4.2	394	12	BE749870	BE749870 200930 MA
519	36	4.2	316	14	BO637699	BO637699 hel3d12.x	592	36	4.2	394	14	BM685815	BM685815 UT-E-CXO-
520	36	4.2	318	10	AM194282	AM194282 xmo08c10.x	593	36	4.2	394	14	BM889036	BM889036 KB04e06.y
521	36	4.2	318	10	BE059034	BE059034 sm24c12.y	594	36	4.2	396	9	AI016656	AI016656 ow96a04.x
522	36	4.2	320	12	BG629409	BG629409 CC-e8E1CL	595	36	4.2	401	9	AI479904	AI479904 tm70g08.x
523	36	4.2	321	9	AU032470	AU032470 AU032470	596	36	4.2	402	9	AI395109	AI395109 MA002437.
524	36	4.2	321	12	BF159593	BF159593 601769091	597	36	4.2	402	9	AI796651	AI796651 AU796651
525	36	4.2	322	13	BI509013	BI509013 BB170009A	598	36	4.2	403	13	BG953673	BG953673 CM4-CT016
526	36	4.2	323	14	AT001084	AT001084 AT001084	599	36	4.2	403	13	BI042924	BI042924 CM4-OT016
527	36	4.2	323	14	BM807723	BM807723 fd17e11.y	600	36	4.2	403	13	BM023352	BM023352 1e79c10.y
528	36	4.2	325	9	AA490112	AA490112 ab02e04.s	601	36	4.2	404	13	BM356983	BM356983 1e79c10.y
529	36	4.2	326	9	AU092544	AU092544 AU092544	602	36	4.2	408	17	AZ928338	AZ928338 479.dif08
530	36	4.2	327	10	AM631107	AM631107 b103h04.x	603	36	4.2	410	9	AI793599	AI793599 C15105
531	36	4.2	328	14	BQ392654	BQ392654 NISC mq25	604	36	4.2	411	13	BM149510	BM149510 L0813E05-
532	36	4.2	328	13	AA764860	AA764860 n261q03.s	605	36	4.2	411	13	BM149510	BM149510 L0813E05-
533	36	4.2	328	13	BI381114	BI381114 361783 MA	606	36	4.2	411	13	BM149510	BM149510 L0813E05-
534	36	4.2	333	13	BG671750	BG671750 DKNBXD05	607	36	4.2	412	9	AI297783	AI297783
535	36	4.2	333	13	BM568122	BM568122 sak99d10.	608	36	4.2	412	9	AI297783	AI297783
536	36	4.2	334	14	BQ100126	BQ100126 ph72g08.y	609	36	4.2	412	13	BO393282	BO393282
537	36	4.2	334	17	AQ676101	AQ676101 HS 5501 B	610	36	4.2	416	9	AA085810	AA085810
538	36	4.2	335	9	AI220836	AI220836 q907c09.x	611	36	4.2	416	13	BM114951	BM114951
539	36	4.2	336	14	BM801080	BM801080 AGENCOURT	612	36	4.2	417	17	CNS02AH7	CNS02AH7
540	36	4.2	338	9	AU261877	AU261877 AU261877	613	36	4.2	418	13	BM030865	BM030865
541	36	4.2	342	14	BQ099940	BQ099940 ph70c08.y	614	36	4.2	421	17	CNS01ZGB	CNS01ZGB
542	36	4.2	343	9	AI662542	AI662542 mc15h06.x	615	36	4.2	424	13	BM030865	BM030865
543	36	4.2	345	9	AI894196	AI894196 mc78c01.x	616	36	4.2	427	10	BE507504	BE507504
544	36	4.2	346	12	BG783849	BG783849 SEAMC003	617	36	4.2	428	13	BM558147	BM558147
545	36	4.2	346	13	BM356276	BM356276 k61h02.y	618	36	4.2	434	9	AJ437804	AJ437804
546	36	4.2	349	14	BQ448961	BQ448961 CXESTY453	619	36	4.2	436	10	AV707923	AV707923
547	36	4.2	350	9	AA815020	AA815020 oa88a09.s	620	36	4.2	437	12	AU263204	AU263204
548	36	4.2	350	9	AI147037	AI147037 qa82c07.x	621	36	4.2	437	12	BE878789	BE878789
549	36	4.2	350	9	AL595503	AL595503 AL595503	622	36	4.2	438	13	BM529056	BM529056
550	36	4.2	352	12	BF904247	BF904247 MR3-MT032	623	36	4.2	442	14	BM099736	BM099736
551	36	4.2	353	13	BI744627	BI744627 rK31c06.y	624	36	4.2	444	14	BO299736	BO299736
552	36	4.2	355	9	AA629768	AA629768 aec5d10.s	625	36	4.2	446	13	BI944215	BI944215
553	36	4.2	356	13	BI424252	BI424252 sah67E01.	626	36	4.2	446	14	RA9284	RA9284
554	36	4.2	356	13	BM565211	BM565211 ih39c01.y	627	36	4.2	447	10	BE038744	BE038744
555	36	4.2	357	12	BG663614	BG663614 DRAMLC07	628	36	4.2	448	13	BC384653	BC384653
556	36	4.2	358	10	AM669118	AM669118 112295 MA	629	36	4.2	448	12	BM134794	BM134794
557	36	4.2	358	12	BG652938	BG652938 sad01h07.	630	36	4.2	449	10	BE442733	BE442733
558	36	4.2	359	9	AI458732	AI458732 tj37e09.x	631	36	4.2	450	9	AA190985	AA190985
559	36	4.2	360	9	AL799082	AL799082 AL799082	632	36	4.2	452	14	BO520771	BO520771
560	36	4.2	360	17	CNS04A3B	AL281360 Tetrarodon	633	36	4.2	453	12	BG662539	BG662539
561	36	4.2	361	9	AL663475	AL663475 AL663475	634	36	4.2	453	13	BM277993	BM277993
562	36	4.2	363	10	AM541819	AM541819 C0149H07-	635	36	4.2	455	12	BM301653	BM301653
563	36	4.2	363	14	C90234	C90234 C90234 Dict	636	36	4.2	456	13	BM125312	BM125312
564	36	4.2	366	9	AL514543	AL514543 AL514543	637	36	4.2	457	13	BO945670	BO945670
565	36	4.2	369	9	AL515389	AL515389 AL515389	638	36	4.2	459	9	AI436380	AI436380
566	36	4.2	370	9	AA014418	AA014418 m166g11.x	639	36	4.2	459	14	RA4267	RA4267
567	36	4.2	371	9	AI653576	AI653576 tq96e05.x	640	36	4.2	460	9	AI181025	AI181025
568	36	4.2	371	9	AJ318316	AJ318316 AJ318316	641	36	4.2	463	9	AL703354	AL703354
569	36	4.2	371	13	BM516033	BM516033 kj72f01.y	642	36	4.2	463	10	AV976393	AV976393
570	36	4.2	374	9	AL044337	AL044337 DKF2p434A	643	36	4.2	464	9	AV976393	AV976393
571	36	4.2	374	10	AV723825	AV723825 AV723825	644	36	4.2	465	9	AJ318285	AJ318285
572	36	4.2	374	14	BQ385824	BQ385824 NISC mm14	645	36	4.2	466	13	BI944967	BI944967
573	36	4.2	375	9	AA708335	AA708335 zg07e01.s	646	36	4.2	468	10	AV717525	AV717525
574	36	4.2	376	9	AI220834	AI220834 q907c07.x	647	36	4.2	468	13	BI534636	BI534636
575	36	4.2	378	10	AM395009	AM395009 bh38c12.y	648	36	4.2	469	9	AU075713	AU075713
576	36	4.2	378	13	BI079131	BI079131 602873422	649	36	4.2	470	9	AU263029	AU263029
577	36	4.2	381	9	AL024005	AL024005 K8224b23	650	36	4.2	470	9	AU263034	AU263034
578	36	4.2	381	9	AL1719429	AL1719429 AL1719429	651	36	4.2	474	13	BI738206	BI738206
579	36	4.2	383	9	AI334191	AI334191 qp92c12.x	652	36	4.2	475	12	BG842644	BG842644
580	36	4.2	383	9	AI334191	AI334191 qp92c12.x	653	36	4.2	476	17	CNS02YXW	CNS02YXW
581	36	4.2	384	9	AI061173	AI061173 an25c09.x	654	36	4.2	477	17	AA550666	AA550666
582	36	4.2	385	12	BI013373	BI013373 602284981	655	36	4.2	479	9	AA761205	AA761205
583	36	4.2	388	13	BI895413	BI895413 ERESTEa38	656	36	4.2	479	9	AT002581	AT002581
584	36	4.2	388	14	BM774308	BM774308 ra91h07.y	657	36	4.2	479	9	AU095274	AU095274
585	36	4.2	390	17	CNS025HD	AL182074 Tetrarodon	658	36	4.2	479	12	BG261785	BG261785
586	36	4.2	391	9	AU263307	AU263307 AU263307	659	36	4.2	480	9	AU077487	AU077487
587	36	4.2	392	9	AL514511	AL514511 AL514511	660	36	4.2	480	9	BG671339	BG671339
588	36	4.2	392	13	BI697571	BI697571 603348738	661	36	4.2	481	12	BF535232	BF535232
589	36	4.2	393	10	AM231676	AM231676 687018D06	662	36	4.2	481	12	BF535232	BF535232
590	36	4.2	394	10	AV693913	AV693913 AV693913	663	36	4.2	482	14	BQ582233	BQ582233

C 664	36	4.2	483	17	AA550010	AA550010 1089m3 gm	C 737	36	4.2	650	10	AM215874	AM215874 up02e02.x
C 665	36	4.2	484	17	AA15461	AA15461 mc58a02.x	C 738	36	4.2	650	14	C23659	C23659 up02e02.x
C 666	36	4.2	484	17	AA550000	AA550000 1079m3 gm	C 739	36	4.2	653	9	AL520584	AL520584 up02e02.x
C 667	36	4.2	487	10	AA229259	AA229259 u008e02.y	C 740	36	4.2	654	9	AL732282	AL732282 up02e02.x
C 668	36	4.2	493	10	BI964855	BI964855 1e57e10.y	C 741	36	4.2	663	10	AM215881	AM215881 up02e01.x
C 669	36	4.2	498	13	AV956535	AV956535 AV956535	C 742	36	4.2	664	13	BI522894	BI522894 up02e01.x
C 670	36	4.2	494	13	BM380421	BM380421 BM380421	C 743	36	4.2	665	17	CNS0421A	CNS0421A up02e01.x
C 671	36	4.2	494	14	BO841769	BO841769 WHE2985.A	C 744	36	4.2	673	11	AG150663	AG150663 Pan trogl
C 672	36	4.2	495	13	BM563321	BM563321 AGENCOURT	C 745	36	4.2	674	11	BC018428	BC018428 Homo sapi
C 673	36	4.2	497	13	BM380954	BM380954 MEST127-G	C 746	36	4.2	683	12	BE245284	BE245284 AV682476
C 674	36	4.2	501	14	BO255537	BO255537 MEST127-G	C 747	36	4.2	686	14	C93771	C93771 up02e01.x
C 675	36	4.2	502	17	AO882556	AO882556 HS_5407.B	C 748	36	4.2	689	17	CNS04AB8	CNS04AB8 up02e01.x
C 676	36	4.2	504	17	AI436864	AI436864 f034d12.x	C 749	36	4.2	695	14	C93772	C93772 up02e01.x
C 677	36	4.2	506	10	AM555184	AM555184 10251B02-	C 750	36	4.2	700	12	EG026194	EG026194 up02e01.x
C 678	36	4.2	513	14	BF781707	BF781707 602104230	C 751	36	4.2	703	13	BE087461	BE087461 up02e01.x
C 679	36	4.2	513	17	R20548	R20548 yf57f08.81	C 752	36	4.2	707	9	AO077474	AO077474 up02e01.x
C 680	36	4.2	517	14	BM511474	BM511474 BCGDN86TF	C 753	36	4.2	707	17	BH567037	BH567037 BCGX056TF
C 681	36	4.2	521	17	CNS022PX	CNS022PX Tetradon	C 754	36	4.2	710	10	BE543359	BE543359 up02e01.x
C 682	36	4.2	521	17	AO262124	AO262124 Tetradon	C 755	36	4.2	711	10	BE543359	BE543359 up02e01.x
C 683	36	4.2	523	14	BO267877	BO267877 NISC FF15	C 756	36	4.2	715	11	AY068531	AY068531 up02e01.x
C 684	36	4.2	523	14	BE878764	BE878764 601493030	C 757	36	4.2	720	12	BG289378	BG289378 up02e01.x
C 685	36	4.2	529	12	BF128553	BF128553 601810801	C 758	36	4.2	721	13	BG289378	BG289378 up02e01.x
C 686	36	4.2	529	14	C99537	C99537 C99537 Rice	C 759	36	4.2	721	13	BG289378	BG289378 up02e01.x
C 687	36	4.2	533	14	BM938298	BM938298 UI-M-BZL-	C 760	36	4.2	724	9	AL721250	AL721250 up02e01.x
C 688	36	4.2	537	9	AL514397	AL514397 AL514397	C 761	36	4.2	725	10	AM332539	AM332539 up02e01.x
C 689	36	4.2	537	13	BI000365	BI000365 MR3-HN006	C 762	36	4.2	725	10	AM332539	AM332539 up02e01.x
C 690	36	4.2	537	13	BI000365	BI000365 MR3-HN006	C 763	36	4.2	736	13	BI489164	BI489164 up02e01.x
C 691	36	4.2	541	14	BO746528	BO746528 UI-M-FAO-	C 764	36	4.2	745	9	AA392077	AA392077 up02e01.x
C 692	36	4.2	541	12	BG623042	BG623042 602647924	C 765	36	4.2	749	17	AO234504	AO234504 up02e01.x
C 693	36	4.2	542	14	BO389705	BO389705 NISC mc09	C 766	36	4.2	750	12	BG21362	BG21362 up02e01.x
C 694	36	4.2	544	13	BM422094	BM422094 V024D01.O	C 767	36	4.2	753	12	BG74234	BG74234 up02e01.x
C 695	36	4.2	545	13	BI654947	BI654947 603283308	C 768	36	4.2	753	12	BF160261	BF160261 up02e01.x
C 696	36	4.2	547	12	BF579015	BF579015 602096134	C 769	36	4.2	756	13	BI529311	BI529311 up02e01.x
C 697	36	4.2	550	9	AO263167	AO263167 AO263167	C 770	36	4.2	756	13	BI529311	BI529311 up02e01.x
C 698	36	4.2	554	14	BM320510	BM320510 NISC n105	C 771	36	4.2	761	12	BG568336	BG568336 up02e01.x
C 699	36	4.2	555	13	BM394602	BM394602 50072-2-4	C 772	36	4.2	765	17	BH148938	BH148938 up02e01.x
C 700	36	4.2	560	14	BO798516	BO798516 EST 685.G	C 773	36	4.2	778	12	BG740883	BG740883 up02e01.x
C 701	36	4.2	560	14	BO799688	BO799688 EST 1857	C 774	36	4.2	778	12	AG058583	AG058583 Pan trogl
C 702	36	4.2	563	14	BO522774	BO522774 NISC n118	C 775	36	4.2	785	17	BG288001	BG288001 up02e01.x
C 703	36	4.2	564	14	BO394823	BO394823 NISC-ng11	C 776	36	4.2	804	13	BI490367	BI490367 up02e01.x
C 704	36	4.2	567	9	AL511470	AL511470 AL511470	C 777	36	4.2	805	13	BG392282	BG392282 up02e01.x
C 705	36	4.2	567	9	AL511470	AL511470 AL511470	C 778	36	4.2	812	17	AG044252	AG044252 up02e01.x
C 706	36	4.2	569	13	BM269209	BM269209 MEST1406-G	C 779	36	4.2	819	13	BM397721	BM397721 up02e01.x
C 707	36	4.2	572	14	C90451	C90451 C90451 Dict	C 780	36	4.2	821	9	AA987044	AA987044 up02e01.x
C 708	36	4.2	574	13	BI523033	BI523033 603175528	C 781	36	4.2	821	10	AV758661	AV758661 up02e01.x
C 709	36	4.2	576	14	BM939047	BM939047 UI-M-BZL-	C 782	36	4.2	821	10	BF383726	BF383726 up02e01.x
C 710	36	4.2	582	11	AY109035	AY109035 Zee mayb	C 783	36	4.2	826	12	BG619471	BG619471 up02e01.x
C 711	36	4.2	584	9	AA075221	AA075221 zm87f03.s	C 784	36	4.2	836	12	BG029912	BG029912 up02e01.x
C 712	36	4.2	585	14	BO523914	BO523914 NISC-no01	C 785	36	4.2	837	9	AI255435	AI255435 up02e01.x
C 713	36	4.2	586	12	BF782515	BF782515 602106531	C 786	36	4.2	847	13	BI734120	BI734120 up02e01.x
C 714	36	4.2	587	14	BO526974	BO526974 NISC no19	C 787	36	4.2	848	17	CNS0257Y	CNS0257Y up02e01.x
C 715	36	4.2	588	14	BO388777	BO388777 NISC-mq03	C 788	36	4.2	858	12	BG670832	BG670832 up02e01.x
C 716	36	4.2	596	9	AO263021	AO263021 AO263021	C 789	36	4.2	858	12	BG670832	BG670832 up02e01.x
C 717	36	4.2	599	14	BO251019	BO251019 TaR25034B	C 790	36	4.2	862	12	BE970652	BE970652 up02e01.x
C 718	36	4.2	601	9	AO262588	AO262588 AO262588	C 791	36	4.2	869	12	BE970652	BE970652 up02e01.x
C 719	36	4.2	602	13	BI686789	BI686789 603391944	C 792	36	4.2	889	13	BI249196	BI249196 up02e01.x
C 720	36	4.2	603	14	C24400	C24400 C24400 Dict	C 793	36	4.2	895	12	BE884698	BE884698 up02e01.x
C 721	36	4.2	604	9	AO261895	AO261895 AO261895	C 794	36	4.2	915	13	BM425365	BM425365 up02e01.x
C 722	36	4.2	606	13	BM079963	BM079963 MEST102-A	C 795	36	4.2	925	17	BI212383	BI212383 up02e01.x
C 723	36	4.2	608	14	BO522946	BO522946 NISC n119	C 796	36	4.2	934	17	BI64537	BI64537 up02e01.x
C 724	36	4.2	610	14	BO739519	BO739519 PFRSTab4	C 797	36	4.2	936	12	BG682316	BG682316 up02e01.x
C 725	36	4.2	613	17	AO922590	AO922590 RPCI-23-2	C 798	36	4.2	939	17	CNS01UPU	CNS01UPU up02e01.x
C 726	36	4.2	614	17	CNS0152H	AI104945 Drosophill	C 799	36	4.2	940	17	CNS046PD	CNS046PD up02e01.x
C 727	36	4.2	615	14	C25773	C25773 C25773 Dict	C 800	36	4.2	950	10	BE039592	BE039592 up02e01.x
C 728	36	4.2	619	12	AL514935	AL514935 AL514935	C 801	36	4.2	957	12	BE731649	BE731649 up02e01.x
C 729	36	4.2	621	12	BE844577	BE844577 EST322.AP	C 802	36	4.2	964	17	BM416130	BM416130 up02e01.x
C 730	36	4.2	622	14	BO524805	BO524805 NISC no07	C 803	36	4.2	978	17	CNS00590	CNS00590 up02e01.x
C 731	36	4.2	623	14	BO391283	BO391283 NISC-ng17	C 804	36	4.2	979	9	AL54968	AL54968 up02e01.x
C 732	36	4.2	626	11	AY068430	AY068430 Schmitdee	C 805	36	4.2	980	9	AL54968	AL54968 up02e01.x
C 733	36	4.2	630	17	BO524081	BO524081 NISC-no02	C 806	36	4.2	983	13	BI489205	BI489205 up02e01.x
C 734	36	4.2	637	17	CNS0245B	AL180344 Tetradon	C 807	36	4.2	989	17	AZ549932	AZ549932 up02e01.x
C 735	36	4.2	643	9	AO262240	AO262240 AO262240	C 808	36	4.2	1012	14	BO939005	BO939005 up02e01.x
C 736	36	4.2	643	9	AO262240	AO262240 AO262240	C 809	36	4.2	1024	12	BF792715	BF792715 up02e01.x



C 810	36	4.2	1027	17	AZ167189	C 883	35	4.1	121	14	H74729	H74729 274 Deletio
C 811	36	4.2	1040	13	B1521877	884	35	4.1	122	14	BQ235651	BQ235651 hdb3n09.g
C 812	36	4.2	1051	9	AL524212	885	35	4.1	125	13	B111806	B111806 602897836
C 813	36	4.2	1059	12	BF183554	886	35	4.1	126	10	AV961814	AV961814 AV961814
C 814	36	4.2	1064	17	CNS02DX	887	35	4.1	126	10	AM426501	AM426501 60745 MAR
C 815	36	4.2	1095	12	BF525514	888	35	4.1	126	13	B1743529	B1743529 kx51d12.y
C 816	36	4.2	1100	11	BC016864	889	35	4.1	129	14	C90008	C90008 C90008 Dict
C 817	36	4.2	1101	13	B1489094	C 890	35	4.1	130	9	AL1370359	AL1370359 McBA37C10
C 818	36	4.2	1101	17	CNS003B4	891	35	4.1	132	9	AA607405	AA607405 vo39f09.x
C 819	36	4.2	1101	17	CNS000JP	892	35	4.1	135	12	BE874706	BE874706 601485832
C 820	36	4.2	1101	17	CNS00Y6	893	35	4.1	137	12	BF672397	BF672397 602150637
C 821	36	4.2	1101	17	CNS0171E	894	35	4.1	137	12	BE976640	BE976640 b554b03.y
C 822	36	4.2	1112	12	BG116078	895	35	4.1	137	13	BE956252	BE956252 QV1 -CT075
C 823	36	4.2	1112	13	BQ952170	896	35	4.1	139	9	AA129387	AA129387 qc34d08.x
C 824	36	4.2	1117	14	BM415411	897	35	4.1	140	9	AI1340633	AI1340633 lb32b08.x
C 825	36	4.2	1145	13	B1521872	898	35	4.1	140	9	AI162232	AI162232 qk09e09.x
C 826	36	4.2	1147	13	B1499110	899	35	4.1	143	9	AA189259	AA189259 mcb9a03.x
C 827	36	4.2	1150	12	BG532790	C 900	35	4.1	143	13	BE988379	BE988379 MR3 -HT120
C 828	36	4.2	1162	12	BE966022	C 901	35	4.1	144	9	AL502847	AL502847 AL502847
C 829	36	4.2	1162	13	B1524383	C 902	35	4.1	144	12	BF751315	BF751315 RC3 -BN042
C 830	36	4.2	1165	13	B1522984	C 903	35	4.1	144	12	BF751315	BF751315 RC3 -BN042
C 831	36	4.2	1175	10	BE421167	C 904	35	4.1	145	10	AM949205	AM949205 QV4 -FT000
C 832	36	4.2	1175	10	BE421167	C 905	35	4.1	146	9	AU072297	AU072297 AU072297
C 833	36	4.2	1203	13	CNS00ZWL	906	35	4.1	146	13	BT941674	BT941674 sc84a11.y
C 834	36	4.2	1203	13	BM466037	907	35	4.1	147	12	BE165052	BE165052 602343436
C 835	36	4.2	1269	14	BE967005	908	35	4.1	147	13	BE991133	BE991133 MR2 -HT118
C 836	36	4.2	1270	13	BG964449	909	35	4.1	148	9	AA289204	AA289204 VB33f01.x
C 837	36	4.2	1375	13	B1523146	C 910	35	4.1	148	12	BE874133	BE874133 601484317
C 838	36	4.2	1392	12	BG036520	C 911	35	4.1	149	9	AI139814	AI139814 c156d06.x
C 839	36	4.2	1434	13	B1524419	C 912	35	4.1	149	13	BT744597	BT744597 rk89a05.y
C 840	36	4.2	1475	13	B1524374	C 913	35	4.1	151	13	BT268634	BT268634 NF014A11G
C 841	36	4.2	1553	13	B1524338	C 914	35	4.1	151	13	BT002433	BT002433 MR3 -HN015
C 842	36	4.2	1563	12	BG166366	C 915	35	4.1	153	12	BE874163	BE874163 601484347
C 843	36	4.2	1613	10	BE616336	C 916	35	4.1	155	10	BE232007	BE232007 1362909 MA
C 844	36	4.2	1657	10	BE964263	C 917	35	4.1	156	9	AA036486	AA036486 mh17f09.x
C 845	36	4.2	1759	12	EG681855	C 918	35	4.1	157	9	AU071392	AU071392 AU071392
C 846	36	4.2	1819	11	BC013681	C 919	35	4.1	160	9	AU175799	AU175799 AU175799
C 847	36	4.2	1996	11	BC013320	C 920	35	4.1	160	13	BT001154	BT001154 MR3 -HN012
C 848	36	4.2	1996	11	BC021044	C 921	35	4.1	161	9	AL386476	AL386476 McBC34H06
C 849	36	4.2	4195	11	BC021044	C 922	35	4.1	161	9	AU095555	AU095555 AU095555
C 850	35	4.1	56	14	BQ520931	C 923	35	4.1	162	9	AL726422	AL726422 AL726422
C 851	35	4.1	57	13	BT743047	C 924	35	4.1	162	13	BT268676	BT268676 NF021H04G
C 852	35	4.1	58	10	AV954391	C 925	35	4.1	162	13	BM076525	BM076525 sak31904.
C 853	35	4.1	58	14	BQ521646	C 926	35	4.1	164	13	BG991137	BG991137 MR2 -HT118
C 854	35	4.1	68	10	AV971432	C 927	35	4.1	164	13	BT095172	BT095172 sae04c06.
C 855	35	4.1	73	10	AV957892	C 928	35	4.1	165	9	AA265842	AA265842 m270g01.x
C 856	35	4.1	76	9	AA572761	C 929	35	4.1	165	12	AM101024	AM101024 sd64f11.y
C 857	35	4.1	93	14	BQ234257	C 930	35	4.1	165	10	BF535331	BF535331 602052106
C 858	35	4.1	93	14	BQ234543	C 931	35	4.1	165	13	BG950562	BG950562 PM0 -CT070
C 859	35	4.1	93	14	BQ235425	C 932	35	4.1	165	13	BG977643	BG977643 MR2 -CT018
C 860	35	4.1	93	14	BQ788232	C 933	35	4.1	166	14	BQ145504	BQ145504 NF009F12G
C 861	35	4.1	94	14	BQ788464	C 934	35	4.1	167	10	BE381058	BE381058 601271405
C 862	35	4.1	100	13	BT024076	C 935	35	4.1	167	14	C99396	C99396 C99396 R1ce
C 863	35	4.1	102	2	HSW003325	C 936	35	4.1	168	10	AM101810	AM101810 sd70g06.y
C 864	35	4.1	103	10	AV948612	C 937	35	4.1	168	13	BE996101	BE996101 MR4 -HT119
C 865	35	4.1	105	10	AA941045	C 938	35	4.1	169	9	AI1932889	AI1932889 w039c05.x
C 866	35	4.1	105	13	BI671965	C 939	35	4.1	169	9	AI1964507	AI1964507 436011H05
C 867	35	4.1	106	2	HSW003326	C 940	35	4.1	169	9	AL037030	AL037030 DFE2p564H
C 868	35	4.1	106	14	C25768	C 941	35	4.1	169	9	AL636282	AL636282 AL636282
C 869	35	4.1	107	12	BG11851	C 942	35	4.1	169	12	BE896769	BE896769 601437442
C 870	35	4.1	107	13	BI492602	C 943	35	4.1	169	13	BG983805	BG983805 MR4 -CN014
C 871	35	4.1	110	12	BF771374	C 944	35	4.1	169	13	BQ835013	BQ835013 Po ad 06B
C 872	35	4.1	112	13	BJ365473	C 945	35	4.1	171	13	BG981076	BG981076 MR3 -CN014
C 873	35	4.1	112	14	BQ826006	C 946	35	4.1	171	14	BM689264	BM689264 kb08a03.y
C 874	35	4.1	113	13	BJ367608	C 947	35	4.1	172	9	AU175838	AU175838 AU175838
C 875	35	4.1	114	9	AI280525	C 948	35	4.1	172	13	BI538078	BI538078 428454 MA
C 876	35	4.1	116	2	AI925163	C 949	35	4.1	173	10	BE482369	BE482369 168131 BA
C 877	35	4.1	116	9	BF766531	C 950	35	4.1	173	13	BI021161	BI021161 CM0 -MT035
C 878	35	4.1	117	14	BQ788368	C 951	35	4.1	173	13	BM620920	BM620920 170006874
C 879	35	4.1	118	14	BQ788245	C 952	35	4.1	174	9	AA504521	AA504521 aa60n12.s
C 880	35	4.1	119	10	AA431018	C 953	35	4.1	174	13	BI002918	BI002918 MR3 -HN015
C 881	35	4.1	121	10	AM655414	C 954	35	4.1	175	9	AI1334890	AI1334890 lb33a07.x
C 882	35	4.1	121	10	AM655414	C 955	35	4.1	175	9	AL043469	AL043469 DFE2p434A

956	35	4.1	175	10	AM119465	AM119465 sc47406.Y
957	35	4.1	175	14	BM896455	BM896455 p151d01.Y
958	35	4.1	175	14	BQ299084	BQ299084 sa054912.Y
959	35	4.1	176	12	BG653001	BG653001 sad43f09.Y
960	35	4.1	176	13	BG981095	BG981095 MR3-CN014
961	35	4.1	176	13	BG991084	BG991084 MR2-HT118
962	35	4.1	178	9	AU058315	AU058315 AU058315
963	35	4.1	179	9	A1874162	A1874162 wmd8911.x
964	35	4.1	179	9	AU074259	AU074259 AU074259
965	35	4.1	179	13	BI042748	BI042748 CM4-OT016
966	35	4.1	180	9	AU175646	AU175646 AU175646
967	35	4.1	181	9	A1567894	A1567894 lre3c08.x
968	35	4.1	181	14	BQ826005	BQ826005 OK-YZ-B54
969	35	4.1	181	14	C93784	C93784 C93784.D1Ct
970	35	4.1	182	12	BG001253	BG001253 IL0-GM028
971	35	4.1	183	10	AM898582	AM898582 RCL-NN007
972	35	4.1	183	12	BG577746	BG577746 Th_ad10G
973	35	4.1	184	10	AV933802	AV933802 AV933802
974	35	4.1	184	10	AM176457	AM176457 OV0-CT001
975	35	4.1	184	12	BG628853	BG628853 CC-estf1cL
976	35	4.1	185	10	AM196138	AM196138 XM05A10.X
977	35	4.1	185	12	BF811784	BF811784 MR2-CI018
978	35	4.1	185	10	BE807445	BE807445 gs82c11.Y
979	35	4.1	185	13	BI014127	BI014127 CM4-ET023
980	35	4.1	185	13	BI703945	BI703945 Kt24b04.Y
981	35	4.1	185	14	BQ145479	BQ145479 NF09D12G
982	35	4.1	186	9	AA796636	AA796636 vp31h08.r
983	35	4.1	187	9	AU074867	AU074867 AU074867
984	35	4.1	187	14	BQ146233	BQ146233 NF013H10G
985	35	4.1	187	14	C93857	C93857 C93857.D1Ct
986	35	4.1	188	10	BE519303	BE519303 945007C09
987	35	4.1	188	13	BI001827	BI001827 MR3-HN014
988	35	4.1	188	17	AZ576647	AZ576647 AST-T14C0
989	35	4.1	189	10	AM164346	AM164346 se71c10.Y
990	35	4.1	189	14	BQ577707	BQ577707 3524_1.41
991	35	4.1	189	14	C92531	C92531 C92531.D1Ct
992	35	4.1	189	14	BM506342	BM506342 1142f01.Y
993	35	4.1	190	13	BI012867	BI012867 1121h07.Y
994	35	4.1	191	13	BG956187	BG956187 OV1-CT075
995	35	4.1	191	13	BI002366	BI002366 MR3-HN015
996	35	4.1	191	13	BI695960	BI695960 603346151
997	35	4.1	191	13	BI701585	BI701585 sail7f01.Y
998	35	4.1	191	13	BI747326	BI747326 tm38f06.Y
999	35	4.1	191	13	BI941625	BI941625 sc80b10.Y
1000	35	4.1	191	13	BI945299	BI945299 sb53d06.Y

## ALIGNMENTS

RESULT 1	562 bp	mRNA	linear	EST 01-MAR-2002
BM732121	562 bp	mRNA	linear	EST 01-MAR-2002
LOCUS	562 bp	mRNA	linear	EST 01-MAR-2002
DEFINITION	562 bp	mRNA	linear	EST 01-MAR-2002
DESCRIPTION	562 bp	mRNA	linear	EST 01-MAR-2002
ACCESSION	562 bp	mRNA	linear	EST 01-MAR-2002
VERSION	562 bp	mRNA	linear	EST 01-MAR-2002
KEYWORDS	562 bp	mRNA	linear	EST 01-MAR-2002
SOURCE	562 bp	mRNA	linear	EST 01-MAR-2002
ORGANISM	562 bp	mRNA	linear	EST 01-MAR-2002
REFERENCE	562 bp	mRNA	linear	EST 01-MAR-2002
AUTHORS	562 bp	mRNA	linear	EST 01-MAR-2002
TITLE	562 bp	mRNA	linear	EST 01-MAR-2002

JOURNAL	Unpublished (1999)
COMMENT	Contact: Shoemaker R./Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.edu This clone is available through: RegGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: cun@reggen.com web site: www.reggen.com Seq primer: -408P from Gibco High quality sequence stop: 421.
FEATURES	source
source	1..562 /organism="Glycine max" /db_xref="taxon:3847" /clone="SOYBEAN CLONE ID: Gm-c1061-4678" /clone_id="Gm-c1061" /tissue_type="mature flowers of field grown plants" /lab_host="DH10B" /note="vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10 host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."
BASE COUNT	187 a 115 c 128 g 132 t
ORIGIN	Query Match 65.1%; Score 562; DB 14; Length 562; Best Local Similarity 100.0%; Pred. No. 3.3e-128; Mismatches 0; Indels 0; Gaps 0; Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
10	CTATCACACACACAAATGATGATTAACAAACAGTGAAGCTCTCAAGTCTCGA 69
1	CTATCACACACACAAATGATGATTAACAAACAGTGAAGCTCTCAAGTCTCGA 60
70	AGTGAAGAAAGGCGCTTGACAAATGAAAGAGCTGATCTGATGAACTATTTGAAA 129
61	AGTGAAGAAAGGCGCTTGACAAATGAAAGAGCTGATCTGATGAACTATTTGAAA 120
130	TCATGGGGAAGGCTGTTGAACTCTTTGGCCAAAGCTGCTGCTCAACGTAACGAAA 189
121	TCATGGGGAAGGCTGTTGAACTCTTTGGCCAAAGCTGCTGCTCAACGTAACGAAA 180
190	GAGTGGCCGCTTAAGTGGCTTAATTAATCTCCGCTGATGTTGAAGAGGAATATTAC 249
181	GAGTGGCCGCTTAAGTGGCTTAATTAATCTCCGCTGATGTTGAAGAGGAATATTAC 240
250	ACCGAGGAACAACTTTTATGATGATGAGCTTCAAGCAAGAGTGGGAAAACAGTGGTCCA 309
241	ACCGAGGAACAACTTTTATGATGATGAGCTTCAAGCAAGAGTGGGAAAACAGTGGTCCA 300
310	AATGGCAAGCATCTACCTGGAAGAGCTGATGATGATGATGATGATGATGATGATGAT 369
301	AATGGCAAGCATCTACCTGGAAGAGCTGATGATGATGATGATGATGATGATGATGAT 360
370	GATTCAGAGCATCATCAAGCAAGCTGGAACCTTTCAGCAACAGAGTATGATATTTCTGA 429
361	GATTCAGAGCATCATCAAGCAAGCTGGAACCTTTCAGCAACAGAGTATGATATTTCTGA 420
430	GATTAAGATCATCAAGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 489
421	GATTAAGATCATCAAGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 480

QY 490 GTATTCACACCTGTTATCAAGAAATGTAGACCAATTTTCAACTGAGTCCCTACAT 549  
 DB 481 GTATTCACACCTGTTATCAAGAAATGTAGACCAATTTTCAACTGAGTCCCTACAT 540  
 QY 550 TAACTCTGATCAATCCAGTTGT 571  
 DB 541 TAACTCTGATCAATCCAGTTGT 562

**RESULT 2**  
 BE820766/c 586 bp mRNA linear EST 24-MAY-2001  
**LOCUS** GM70001220H11 Gm-r1070 Glycine max cDNA clone Gm-r1070-4966 3'  
**DEFINITION** mRNA sequence.  
**ACCESSION** BE820766  
**VERSION** BE820766.1 GI:10253000  
**KEYWORDS** EST.  
**SOURCE** soybean.  
**ORGANISM** Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 586)  
 Vodka, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,  
 Erpelid, J., Rapp, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.  
 A Functional Genomics Program for Soybean (NSF 9872565)  
 Unpublished (1999)  
 Other ESTs: AM459279 corresponding to Gm-cl016-5560 (5')  
 Contact: Vodka, L.O., PI, A Functional Genomics Program for  
 Soybean (NSF 9872565)  
 Lewin, H. A., Director, Keck Center for Comparative and Functional  
 Genomics  
 University of Illinois  
 Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA  
 Tel: (217) 244-6147  
 Fax: (217) 333-4582  
 Email: l-vodka@uiuc.edu  
 This clone is available through: Genome Systems, Inc. 4633 World  
 Parkway Circle St. Louis, Missouri 63134. For further information  
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)  
 427-3324 or contact: clones@genomesystems.com or info@genome  
 systems.com web site: www.genomesystems.com  
 Seq primer: 5'-TTTTTTTTTTTTTTTTTTT (A/C/G)-3'.  
**FEATURES**  
**SOURCE**  
 1..586  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="Gm-r1070-4966"  
 /clone\_1ib="Gm-r1070"  
 /note="The library Gm-r1070 is a sequence-driven, reracked  
 set of 9,216 clones selected from cDNA libraries from  
 various tissues and stages of development of soybean that  
 represent 2,639 sequences from immature cotyledons, 1,770  
 from immature seed coats, 3,938 from flowers, and 869  
 from young pods. The 5' ESTs of the source clones from  
 the different libraries was used to select singletons, or  
 a representative of each contig, which were reracked to  
 form library Gm-r1070. The cDNA clones of the reracked  
 Gm-r1070 library were then sequenced at the 3' end. The  
 contig analysis to select unique genes was performed by  
 the laboratory of Ernest Retzel, Center for Computational  
 Genomics and Bioinformatics, University of Minnesota,  
 http://www.cbc.umn.edu/ResearchProjects/soybean/index.html  
 . Reracking was performed by Genome Systems, St. Louis,  
 http://www.genomesystems.com, and 3' sequencing by the  
 Keck Center for Comparative and Functional Genomics,  
 University of Illinois,  
 http://www.life.uiuc.edu/biotech/keck.html. Note: The  
 corresponding 5' EST from each clone in the Gm-r1070  
 library is listed in the 'OTHER EST' field. The detailed  
 information on the source library for each clone can also  
 be obtained by referring to the Genome Systems clone ID of

BASE COUNT 173 a 110 c 100 g 194 t 9 others  
 ORIGIN  
 Query Match 42.9%; Score 370; DB 12; Length 586;  
 Best Local Similarity 99.6%; Pred. No. 1.9e-81;  
 Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 326 CCTGAAAGACATGATATATGATCAAGAACTATTTGGAGGACAGAGTCAGAGACATC 385  
 DB 583 CCTGAAAGACATGATATATGATCAAGAACTATTTGGAGGACAGAGTCAGAGACATC 524  
 QY 386 AACCAAGCTGAGAACTTTGACGAAACAGAGTAGTAATATTTGAGTAATATGACCA 445  
 DB 523 AACCAAGCTGAGAACTTTGACGAAACAGAGTAGTAATATTTGAGTAATATGACCA 464  
 QY 446 GCTAGCACTAGCAATGTTTCCACCATGCTGAGCCCATGAGATATATCTCACCCCT 505  
 DB 463 GCTAGCACTAGCAATGTTTCCACCATGCTGAGCCCATGAGATATATCTCACCCCT 404  
 QY 506 TATCAAGAAATGTTAGAGCATTTTCACTGAGTCCCTACATTTATCTGATCAATCC 565  
 DB 403 TATCAAGAAATGTTAGAGCATTTTCACTGAGTCCCTACATTTATCTGATCAATCC 344  
 QY 566 AGTTGTTGATCAATGACAAACAACAACTTAATCTATTGGAGCATGAGATAGTGTCA 625  
 DB 343 AGTTGTTGATCAATGACAAACAACAACTTAATCTATTGGAGCATGAGATAGTGTCA 284  
 QY 626 ATGCATTTACTGAAACGGTGATTAATATTTATCAAGATAAAGCTAAGTTTGAAGTTCA 685  
 DB 283 ATGCATTTACTGAAACGGTGATTAATATTTATCAAGATAAAGCTAAGTTTGAAGTTCA 224  
 QY 686 TAAGGCTGGAATGCTGTGATTAATAACATATATTGGGTTGTTATATTAAGTAGTTG 745  
 DB 223 TAAGGCTGGAATGCTGTGATTAATAACATATATTGGGTTGTTATATTAAGTAGTTG 164  
 QY 746 ATGTTTGTTTGGTACATTAATTAAGTATGCTGTAATATATGAGAT 797  
 DB 163 ATGTTTGTTTGGTACATTAATTAAGTATGCTGTAATATATGAGAT 112

**RESULT 3**  
 AM459279 431 bp mRNA linear EST 03-DEC-2001  
**LOCUS** SH22H08.Y1 Gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
**DEFINITION** Gm-cl016-5560 5' similar to TR:Q9ZTD5 Q9ZTD5 PUTATIVE TRANSCRIPTION  
 FACTOR ;, mRNA sequence.  
**ACCESSION** AM459279  
**VERSION** AM459279.1 GI:7029496  
**KEYWORDS** EST.  
**SOURCE** soybean.  
**ORGANISM** Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 431)  
 Shoemaker, R., Keim, P., Vodka, L., Erpelid, J., Coryell, V., Khanna,  
 A., Bolla, B., Matra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
 Wile, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers,  
 Y., Peterson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk,  
 R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann,  
 R., Waterson, R. and Wilson, R.  
 Public Soybean EST project  
 Unpublished (1999)  
 Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu

This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com

Insert Length: 723 Std Error: 0.00  
Seq primer: -40RP from Gibco  
High quality sequence stop: 411.

## FEATURES

SOURCE:

1..431  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1016-5560"  
/tissue\_type="Immature flowers of field grown plants"  
/lab\_host="XLI10-Gold"  
/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XLI10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelidg."

BASE COUNT 145 a 81 c 88 g 117 t  
ORIGIN

## Query Match

Best Local Similarity 100.0%; Score 349; DB 10; Length 431;  
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 CCTGGAAGACTGATTAAGATCAAGAACTATTGAGAGACAGATCCGAGACATC 385  
DB 8 CTGGAAGACTGATTAAGATCAAGAACTATTGAGAGACAGATCCGAGACATC 67  
QY 386 AAGCAAGCTGGAACCTTACGACCAACAGATGATTAATTTCTGATTAATGATCACC 445  
DB 68 AAGCAAGCTGGAACCTTACGACCAACAGATGATTAATTTCTGATTAATGATCACC 127  
QY 446 GCTAGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGAGATGATTTCTCCACCTGT 505  
DB 128 GCTAGCACTAGCCATGTTTCCACCATGGCTGAGCCCATGAGATGATTTCTCCACCTGT 187  
QY 506 TATCAAGGATGTTAGAGCCATTTTCAACTGATCCCTCAATTAATCTGATCAATCC 565  
DB 188 TATCAAGGATGTTAGAGCCATTTTCAACTGATCCCTCAATTAATCTGATCAATCC 247  
QY 566 AATTGTTGACCAATGACCAACACATTAATTAATGAGCATGAGGATAGCTGCTCA 625  
DB 248 AATTGTTGACCAATGACCAACACATTAATTAATGAGCATGAGGATAGCTGCTCA 307  
QY 626 ATGCAATTAATGCAACGGTGATTAATTTATCAAGATTAATTAATTAATTAATTA 674  
DB 308 ATGCAATTAATGCAACGGTGATTAATTTATCAAGATTAATTAATTAATTAATTA 356

## RESULT 4

BE058947

LOCUS

DEFINITION

BE058947 450 bp mRNA linear EST 03-DEC-2001  
Gm-c1016-12290 5', similar to TR:Q92TD5 Q92TD5 PUTATIVE  
TRANSCRIPTION FACTOR ;, mRNA sequence.

ACCESSION

BE058947

KEYWORDS

SOURCE

ORGANISM

soybean.  
Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

## REFERENCE

AUTHORS

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 450)  
Shoemaker R., Kelm P., Vodkin L., Erpelidg J., Coryell V., Khanna  
A., Bolla B., Marra M., Hillier L., Kucaba T., Martin J., Beck C.,  
Wyle T., Underwood K., Stepien M., Theising B., Allen M., Bowers  
Y., Person B., Swaller T., Gibbons M., Pape D., Harvey N., Schurk  
R., Ritter E., Kohn S., Shin T., Jackson Y., Cardenas M., McCann  
R., Materson R. and Wilson R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R./Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: ccu@resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 411.

## FEATURES

SOURCE

1..450  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1016-12290"  
/clone\_id="Gm-c1016"  
/tissue\_type="Immature flowers of field grown plants"  
/lab\_host="XLI10-Gold"  
/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XLI10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelidg."

BASE COUNT 151 a 84 c 90 g 125 t  
ORIGIN

## Query Match

Best Local Similarity 99.7%; Score 285; DB 10; Length 450;  
Matches 335; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 339 ATATGAGATCAAGACTATTGAGAGACAGATCCGAGACATCAAGACCTGAGA 398  
DB 36 ATATGAGATCAAGACTATTGAGAGACAGATCCGAGACATCAAGACCTGAGA 95  
QY 339 ACTTTCAGCAACAGATGATTAATTTCTGATTAATTAATTAATTAATTAATTA 456  
DB 96 ACTTTCAGCAACAGATGATTAATTTCTGATTAATTAATTAATTAATTAATTA 155  
QY 459 ATGTTTCCACCATGGCTGAGCCCATGAGATGATTTCTCCACCTGTTTACAGAAATG 518  
DB 156 ATGTTTCCACCATGGCTGAGCCCATGAGATGATTTCTCCACCTGTTTACAGAAATG 215  
QY 519 TAGAGCCATTTTCAACTGATTTCCCTCAATTAATCTGATCAATCAAGTTGTTGACCA 578  
DB 216 TAGAGCCATTTTCAACTGATTTCCCTCAATTAATCTGATCAATCAAGTTGTTGACCA 275  
QY 579 ATGACAAACAACATTAATTAATTTGAGCATGAGGATAGCTGCTCAATCAATTAATCA 638  
DB 276 ATGACAAACAACATTAATTAATTTGAGCATGAGGATAGCTGCTCAATCAATTAATCA 335  
QY 639 ACGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 674

|||||

Db 336 ACGGATTAATATATATACAGATAAACTAAGT 371

RESULT 5  
BE804790 409 bp mRNA linear EST 06-DEC-2001  
LOCUS sale3906.y1 Gm-cl061 glycine max cDNA clone SOYBEAN CLONE ID:  
DEFINITION Gm-cl061-1702 5' similar to TR:Q39028 Q39028 ATMYB2. ; mRNA  
sequence.  
ACCESSION BE804790  
VERSION BE804790  
KEYWORDS GI:10235902  
SOURCE soybean.  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 409)  
REFERENCE Shoemaker,R., Keim,P., Vodkin,L., Erpelidng,J., Coryell,V., Khanna  
AUTHORS 'A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers  
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk  
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann  
,R., Waterston,R. and Wilson,R.  
Public Soybean EST Project  
TITLE Unpublished (1999)  
JOURNAL Contact: Shoemaker R/Public Soybean EST Project  
COMMENT Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: ccu@resgen.com  
Insert Length: 1068 Std Error: 0.00.  
Location/Qualifiers  
FEATURES  
source  
1. 409  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl061-1702"  
/clone\_1lb="Gm-cl061"  
/tissue\_type="mature flowers of field grown plants"  
/lab\_host="DH10B"  
/note="Vector: pBluescript II SK+; Site\_1: EcoRI; Site\_2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from mature flowers of field grown plants for the cultivar  
Raiden. Complementary DNA was synthesized from mRNA using  
a primer consisting of a poly(dT) sequence with a XhoI  
restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the pBluescript vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (GibcoBRL). This library was constructed in the  
laboratory of Dr. Randy Shoemaker."

BASE COUNT 136 a 87 c 97 g 89 t

ORIGIN

Query Match 6.5%; Score 56; DB 12; Length 409;  
Best Local Similarity 100.0%; Pred. No. 8.6e-05;  
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 215 TACCTCGCTCGTGTAGAGAGGGAATATTACACCCGAGGAACAATTGGAT 270  
|||||

Db 220 TACCTCGCTCGTGTAGAGAGGGAATATTACACCCGAGGAACAATTGGAT 275  
|||||

RESULT 6  
BM527606

LOCUS BM527606 501 bp mRNA linear EST 19-FEB-2002  
DEFINITION sale3906.y1 Gm-cl061 glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-cl061-3780 5' similar to TR:Q39028 Q39028 ATMYB2. ; mRNA  
sequence.  
ACCESSION BM527606  
VERSION BM527606  
KEYWORDS GI:18733434  
SOURCE EST.  
ORGANISM soybean.  
Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 501)  
REFERENCE Shoemaker,R., Keim,P., Vodkin,L., Erpelidng,J., Coryell,V., Khanna  
AUTHORS 'A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers  
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk  
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann  
,R., Waterston,R. and Wilson,R.  
Public Soybean EST Project  
TITLE Unpublished (1999)  
JOURNAL Contact: Shoemaker R/Public Soybean EST Project  
COMMENT Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact: ccu@resgen.com web site:  
www.resgen.com  
Seg primer: -40RP from Gibco  
High quality sequence stop: 421.  
Location/Qualifiers  
FEATURES  
source  
1. 501  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-cl061-3780"  
/clone\_1lb="Gm-cl061"  
/tissue\_type="mature flowers of field grown plants"  
/lab\_host="DH10B"  
/note="Vector: pBluescript II SK+; Site\_1: EcoRI; Site\_2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from mature flowers of field grown plants for the cultivar  
Raiden. Complementary DNA was synthesized from mRNA using  
a primer consisting of a poly(dT) sequence with a XhoI  
restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the pBluescript vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (GibcoBRL). This library was constructed in the  
laboratory of Dr. Randy Shoemaker."

BASE COUNT 170 a 104 c 114 g 113 t

ORIGIN

Query Match 6.5%; Score 56; DB 13; Length 501;  
Best Local Similarity 100.0%; Pred. No. 7.2e-05;  
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 215 TACCTCGCTCGTGTAGAGAGGGAATATTACACCCGAGGAACAATTGGAT 270  
|||||

Db 196 TACCTCGCTCGTGTAGAGAGGGAATATTACACCCGAGGAACAATTGGAT 251  
|||||

RESULT 7  
A1930997 523 bp mRNA linear EST 30-NOV-2001  
LOCUS A1930997  
DEFINITION sD45H07.y1 Gm-cl015 glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
Gm-cl015-278 5' similar to TR:Q39028 Q39028 ATMYB2. ; mRNA  
sequence.

ACCESSION A1930997  
 VERSION A1930997.1 GI:566961  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

# REFERENCE

## AUTHORS

1 (bases 1 to 523)  
 Shoemaker, R., Kaim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepcoe, M., Theising, B., Allen, M., Bowers, J., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu

## TITLE

## JOURNAL

## COMMENT

This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com  
 Seq primer: -40RP from Glibco  
 High quality sequence stop: 422.  
 Location/Qualifiers

## FEATURES

### source

1..523  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1015-278"  
 /clone\_id="Gm-c1015"  
 /tissue\_type="Mature flowers, field grown plants"  
 /lab\_host="XLI0-Gold"  
 /note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from mature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR cDNA library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XLI0-gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."

BASE COUNT 172 a 124 c 111 g 115 t 1 others  
 ORIGIN

Query Match 6.5%; Score 56; DB 9; Length 523;  
 Best Local Similarity 100.0%; Pred. No. 7e-05;  
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 215 TACCTCCGTCCTGATGTTAGAGGAGATATTACACCCGAGGAACAATTGGAT 270  
 |||||  
 Db 244 TACCTCCGTCCTGATGTTAGAGGAGATATTACACCCGAGGAACAATTGGAT 299

RESULT 8 552 bp mRNA linear EST 19-FEB-2002  
 LOCUS BMS27774  
 DEFINITION salis903.y1 Gm-c1061 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1061-4134 5' similar to TR.O49020 O49020 MYB-LIKE DNA-BINDING DOMAIN PROTEIN.; mRNA sequence.

ACCESSION BMS27774  
 VERSION BMS27774.1 GI:18733722  
 KEYWORDS EST.

## SOURCE

soybean.  
 Glycine max

## REFERENCE

1 (bases 1 to 552)  
 Shoemaker, R., Kaim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepcoe, M., Theising, B., Allen, M., Bowers, J., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu

## TITLE

This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com  
 Seq primer: -40RP from Glibco  
 High quality sequence stop: 426.  
 Location/Qualifiers

## FEATURES

### source

1..552  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="SOYBEAN CLONE ID: Gm-c1061-4134"  
 /clone\_id="Gm-c1061"  
 /tissue\_type="Mature flowers of field grown plants"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 184 a 118 c 111 g 139 t  
 ORIGIN

Query Match 6.5%; Score 56; DB 13; Length 552;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-05;  
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 215 TACCTCCGTCCTGATGTTAGAGGAGATATTACACCCGAGGAACAATTGGAT 270  
 |||||  
 Db 43 TACCTCCGTCCTGATGTTAGAGGAGATATTACACCCGAGGAACAATTGGAT 98

RESULT 9 782 bp mRNA linear EST 24-MAY-2001  
 LOCUS BMS58316/c  
 DEFINITION GM700005B10E4 Gm-rt1070 Glycine max cDNA clone Gm-rt1070-1759 3', mRNA sequence.

ACCESSION BMS58316  
 VERSION BMS58316.1 GI:9984208  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

glycine.

1 (bases 1 to 782)

REFERENCE  
AUTHORS  
Vodkin,L., Kern,P., Shoemaker,R., Retzel,E., Khanna,A., Corryell,V.,  
Expelding,U., Rapp,C., Shoop,E., Pardinas,J., Liu,L. and Lewin,H.  
A Functional Genomics Program for Soybean (NSF 9872565)  
Unpublished (1999)

TITLE  
JOURNAL  
Other ESTs: A1930997 corresponding to Gm-c1015-278 (5')  
Contact: Vodkin, L.O., PI, A Functional Genomics Program for  
Soybean (NSF 9872565)  
Lewin, H. A., Director, Keck Center for Comparative and Functional  
Genomics  
University of Illinois  
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA  
Tel: (217) 244-6147  
Fax: (217) 333-4582  
Email: l-vodkin@uiuc.edu

This clone is available through: Genome Systems, Inc. 4633 World  
Parkway Circle St. Louis, Missouri 63134. For further information  
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)  
427-3324 or contact: clones@genomesystems.com or info@genome  
systems.com web site: www.genomesystems.com  
Seq primer: 5'-TTTTTTTTTTTTTTTTTTT(A/C/G)-3'.

FEATURES  
source  
Location/Qualifiers  
1..782  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="Gm-r1070-1759"  
/clone\_lib="Gm-r1070"  
/note="The library Gm-r1070 is a sequence-driven, rerecked  
set of 9,216 clones selected from cDNA libraries from  
various tissues and stages of development of soybean that  
represent 2,639 sequences from immature cotyledons, 1,770  
from immature seed coats, 3,938 from flowers, and 869  
from young pods. The 5' ESTs of the source clones from  
the different libraries was used to select singletons, or  
a representative of each contig, which were rerecked to  
form library Gm-r1070. The cDNA clones of the rerecked  
Gm-r1070 library were then sequenced at the 3' end. The  
contig analysis to select unique genes was performed by  
the laboratory of Ernest Retzel, Center for Computational  
Genomics and Bioinformatics, University of Minnesota,  
http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html  
. Rerecking was performed by Genome Systems, St. Louis,  
http://www.genomesystems.com, and 3' sequencing by the  
Keck Center for Comparative and Functional Genomics,  
University of Illinois,  
http://www.life.uiuc.edu/biotech/keck.html. Note: The  
corresponding 5' EST from each clone in the Gm-r1070  
library is listed in the 'OTHER EST' field. The detailed  
information on the source library for each clone can also  
be obtained by referring to the Genome Systems clone ID of  
the original cDNA library that is also listed under  
'OTHER EST'."

BASE COUNT 215 a 144 c 143 g 253 t 27 others

ORIGIN  
Query Match 5.8%; Score 50; DB 10; Length 782;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 CACGCAAACTGGGGAAACAGTGTCTCAAAATTGCCAGCATCTACCTGG 330  
|||||  
Db 612 CACGCAAACTGGGGAAACAGTGTCTCAAAATTGCCAGCATCTACCTGG 563  
|||||

RESULT 10  
AM432229 213 bp mRNA linear EST 03-DEC-2001  
LOCUS bh70908.y1 Gm-c1015 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
DEFINITION Gm-c1015-5175 5', mRNA sequence.  
ACCESSION AM432229  
VERSION AM432229.1 GI:6963536  
KEYWORDS EST.

soybean.

glycine max

1 (bases 1 to 213)

REFERENCE  
AUTHORS  
Shoemaker,R., Kaim,P., Vodkin,L., Expelding,U., Corryell,V., Khanna  
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers  
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk  
R., Ratter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann  
R., Waterston,R. and Wilson,R.  
Public Soybean EST Project  
Unpublished (1999)

TITLE  
JOURNAL  
Public Soybean EST Project  
Contact: Shoemaker R/Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: ccu@resgen.com  
Seq primer: -40RP from Gibco.

FEATURES  
source  
Location/Qualifiers  
1..213  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1015-5175"  
/clone\_lib="Gm-c1015"  
/tissue\_type="Mature flowers, field grown plants"  
/lab\_host="X110-Gold"  
/note="Vector: pBluescript II XR; Site 1: EcoRI, Site 2:  
XhoI. This cDNA library was constructed from mRNA isolated  
from mature flowers of field grown plants. The cDNA  
library was prepared using the Stratagene plasmid kit II  
XR cDNA library construction kit. Complementary DNA was  
synthesized from mRNA using a primer consisting of a poly  
(dT) sequence with a XhoI restriction site. EcoRI adapters  
were ligated to the blunt-ended cDNA fragments followed by  
XhoI digestion. The cDNA fragments were directionally  
cloned into the EcoRI-XhoI restriction site of the  
plasmid vector. The ligated cDNA fragments were  
transformed into X110-Gold host cells. This library was  
constructed by Dr. Randy Shoemaker and Dr. John  
Expelding."

BASE COUNT 69 a 47 c 36 g 61 t

ORIGIN  
Query Match 5.7%; Score 49; DB 10; Length 213;  
Best Local Similarity 100.0%; Pred. No. 0.0075;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 545 ACAATTATCTCTGATCAATCCAGTTGTTGACCAATGACAACAACAACA 593  
|||||  
Db 84 ACAATTATCTCTGATCAATCCAGTTGTTGACCAATGACAACAACAACA 132  
|||||

RESULT 11  
BG652172/c 516 bp mRNA linear EST 29-NOV-2001  
LOCUS bad75b05.y1 Gm-c1051 Glycine max cDNA clone GENOME SYSTEMS CLONE  
DEFINITION ID: Gm-c1051-5842 5' similar to TR:Q9SBG4 Q9SBG4 PUTATIVE  
TRANSCRIPTION FACTOR ;, mRNA sequence.  
ACCESSION BG652172  
VERSION BG652172.1 GI:13789581  
KEYWORDS soybean.  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

REFERENCE  
AUTHORS

Rosidae; eurousids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

1 (bases 1 to 516)

Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Corryell, V., Khanna  
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk  
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)

TITLE  
JOURNAL  
COMMENT

Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@watson.wustl.edu  
Possible reversed clone: similarity on wrong strand This clone is  
available through: ResGen, Invitrogen Corp. 2130 South Memorial  
Parkway Huntville, AL 35801 For further information call: (800  
)-533-4363 or contact via email: ccu@resgen.com  
High quality sequence stop: 310.

## FEATURES

## source:

Location/Qualifiers

1..516

/organism="Glycine max"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1051-5842"

/clone\_1ib="Gm-c1051"

/tissue\_type="floral meristematic mRNA"

/lab\_host="DH10B"

/note="Vector: Bluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI; The cDNA library was constructed from floral  
meristematic mRNA provided by Dr. Halina Knap of Clemson  
University. Complementary DNA was synthesized from mRNA  
using a primer consisting of a poly(dT) sequence with a  
XhoI restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the Bluescript vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (GibcoBRL). This library was constructed in the  
laboratory of Dr. Randy Shoemaker."

BASE COUNT  
ORIGIN

162 a 78 c 89 g 187 t

## Query Match

Best Local Similarity 5.7%; Score 49; DB 12; Length 516;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 545 ACAATTATCCTGATCAATCCAGTTGTTGACCAATGACAACAACA 593

DB 396 ACAATTATCCTGATCAATCCAGTTGTTGACCAATGACAACAACA 348

## RESULT 12

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## glycine

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurousids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

## 1 (bases 1 to 567)

Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Corryell, V., Khanna  
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk  
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)

## Contact: Shoemaker R/Public Soybean EST Project

## Public Soybean EST Project

## Washington University School of Medicine

## 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

## Tel: 314 286 1800

## Fax: 314 286 1810

## Email: est@watson.wustl.edu

## Possible reversed clone: similarity on wrong strand

## This clone is available through: ResGen, Invitrogen Corp. 2130

## South Memorial Parkway Huntville, AL 35801 For further information

## call: (800)-533-4363 or contact: ccu@resgen.com web site:

## www.resgen.com

TITLE  
JOURNAL  
COMMENT

Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk  
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@watson.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800)-533-4363 or contact: ccu@resgen.com web site:  
www.resgen.com

Possible reversed clone: similarity on wrong strand  
Seq primer: -40RP from Gldco  
High quality sequence stop: 422.

## FEATURES

## source:

Location/Qualifiers

1..567

/organism="Glycine max"

/db\_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-c1061-4914"

/clone\_1ib="Gm-c1061"

/tissue\_type="mature flowers of field grown plants"

/lab\_host="DH10B"

/note="Vector: Bluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from mature flowers of field grown plants for the cultivar  
Raden. Complementary DNA was synthesized from mRNA using  
a primer consisting of a poly(dT) sequence with a XhoI  
restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the Bluescript vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (GibcoBRL). This library was constructed in the  
laboratory of Dr. Randy Shoemaker."

BASE COUNT  
ORIGIN

175 a 88 c 101 g 203 t

Query Match 5.7%; Score 49; DB 14; Length 567;  
Best Local Similarity 100.0%; Pred. No. 0.0033;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 545 ACAATTATCCTGATCAATCCAGTTGTTGACCAATGACAACAACA 593

DB 387 ACAATTATCCTGATCAATCCAGTTGTTGACCAATGACAACAACA 339

## RESULT 13

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## glycine

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurousids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

## 1 (bases 1 to 395)

Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Corryell, V., Khanna  
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk  
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)

## Contact: Shoemaker R/Public Soybean EST Project

## Public Soybean EST Project

## Washington University School of Medicine

## 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

## Tel: 314 286 1800

## Fax: 314 286 1810

## Email: est@watson.wustl.edu

## Possible reversed clone: similarity on wrong strand

## This clone is available through: ResGen, Invitrogen Corp. 2130

## South Memorial Parkway Huntville, AL 35801 For further information

## call: (800)-533-4363 or contact: ccu@resgen.com web site:

## www.resgen.com



TITLE  
JOURNAL  
COMMENT

R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: c@resgen.com  
High quality sequence stop: 367.

FEATURES  
source

location/Qualifiers  
1..395

/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl015-877"  
/clone\_1lb="Gm-cl061"  
/tissue\_type="mature flowers of field grown plants"  
/lab\_host="DH10B"  
/note="vector: pBluescript II SK+, Site 1: EcoRI; Site 2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from mature flowers of field grown plants for the cultivar  
Raiden. Complementary DNA was synthesized from mRNA using  
a primer consisting of a poly(dT) sequence with a XhoI  
restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the pBluescript vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (GibcoBRL). This library was constructed in the  
laboratory of Dr. Randy Shoemaker."

BASE COUNT  
ORIGIN

127 a 86 c 103 g 79 t

Query Match  
Best Local Similarity 100.0%; Score 47; DB 12; Length 395;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 215 TACTCCGTCCTGATGTTAGAGAGGATATTACACCCGAGAAC 261  
Db 195 TACTCCGTCCTGATGTTAGAGAGGATATTACACCCGAGAAC 241

RESULT 14  
AM423958  
LOCUS  
DEFINITION  
AM423958 425 bp mRNA linear EST 03-DEC-2001  
5858605.Y1 Gm-cl015 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
Gm-cl015-3993 5' similar to TR:O49020 O49020 MYB-LIKE DNA-BINDING  
DOMAIN PROTEIN.; mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AM423958  
AM423958.1 GI:6951890  
EST.  
soybean.  
Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 425)

REFERENCE  
AUTHORS

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna  
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk  
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
R., Waterston, R. and Wilson, R.

TITLE  
JOURNAL  
COMMENT  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine

Washington University School of Medicine

FEATURES  
source

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: c@resgen.com  
Insert Length: 1178 Std Error: 0.00  
Seq primer: -40RP from Gibco  
High quality sequence stop: 411.

location/Qualifiers  
1..425

/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl015-3993"  
/clone\_1lb="Gm-cl015"  
/tissue\_type="Mature flowers, field grown plants"  
/lab\_host="XL10-Gold"  
/note="vector: pBluescript II XR; Site 1: EcoRI; Site 2:  
XhoI; This cDNA library was constructed from mRNA isolated  
from mature flowers of field grown plants. The cDNA  
library was prepared using the Stratagene pBluescript II  
XR cDNA library construction kit. Complementary DNA was  
synthesized from mRNA using a primer consisting of a poly  
(dT) sequence with a XhoI restriction site. EcoRI adapters  
were ligated to the blunt-ended cDNA fragments followed by  
XhoI digestion. The cDNA fragments were directionally  
cloned into the EcoRI-XhoI restriction site of the  
pBluescript vector. The ligated cDNA fragments were  
transformed into XL10-Gold host cells. This library was  
constructed by Dr. Randy Shoemaker and Dr. John  
Erpelting."

BASE COUNT  
ORIGIN

130 a 95 c 107 g 93 t

Query Match  
Best Local Similarity 100.0%; Score 47; DB 10; Length 425;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 215 TACTCCGTCCTGATGTTAGAGAGGATATTACACCCGAGAAC 261  
Db 160 TACTCCGTCCTGATGTTAGAGAGGATATTACACCCGAGAAC 206

RESULT 15  
BE807621  
LOCUS  
DEFINITION  
BE807621 467 bp mRNA linear EST 06-DEC-2001  
5828605.Y1 Gm-cl061 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
Gm-cl061-33 5' similar to TR:Q39028 Q39028 ATMYB2.; mRNA  
sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BE807621  
BE807621.1 GI:10238733  
EST.  
soybean.  
Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 467)

REFERENCE  
AUTHORS

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna  
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk  
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
R., Waterston, R. and Wilson, R.

TITLE  
JOURNAL  
COMMENT  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800

Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: cu@resgen.com  
 Insert Length: 1403 Std Error: 0.00  
 High quality sequence stop: 411.

## FEATURES

## Source

Location/Qualifiers  
 1..467  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1061-33"  
 /clone\_id="Gm-c1061"  
 /tissue\_type="mature flowers of field grown plants"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
 XhoI; The cDNA library was constructed from mRNA isolated  
 from mature flowers of field grown plants for the cultivar  
 Ralston. Complementary DNA was synthesized from mRNA using  
 a primer consisting of a poly(dT) sequence with a XhoI  
 restriction site. EcoRI adapters were ligated to the  
 blunt-ended cDNA fragments followed by XhoI digestion. The  
 cDNA fragments were directionally cloned into the  
 EcoRI-XhoI restriction site of the pBluescript vector. The  
 ligated cDNA fragments were transformed into DH10B host  
 cells (GibcoBRL). This library was constructed in the  
 laboratory of Dr. Randy Shoemaker."

BASE COUNT 154 a 78 c 105 g 130 t  
 ORIGIN

## Query Match

Best Local Similarity 5.4%; Score 47; DB 12; Length 467;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 215 TACCTCCGTCCTGATGTTAGAAGAGGAATATTACCCGAGGAACA 261  
 ||||||||||||||||||||||||||||||||||||||||||||  
 Db 177 TACCTCCGTCCTGATGTTAGAAGAGGAATATTACCCGAGGAACA 223

Search completed: February 8, 2003, 11:07:15  
 Job time : 2058 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 8, 2003, 10:33:30 ; Search time 2673 Seconds  
(without alignments)  
2242.865 Million cell updates/sec

Title: US-10-021-811-36  
Perfect score: 1138  
Sequence: 1 MDKQCKTSQDPEVRKGPW.....NNINYSMEDSMQMLNGD 206

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODL=frame+ p2n.model -DEV=xjh  
-O=/cgm2\_1/USPTO.epool/US10021811/runat\_03022003\_111222\_25637/app\_query.fasta\_1.391  
-DB=genmb1 -QWMT=fastap -SUFFIX=igc -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=Dlosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTMT=ptio -NORM=ext -HEAPSIZE=570 -MINLEN=0 -MAXLEN=200000000  
-USER=US10021811 @CGN 1.1 2425 @runat\_03022003\_111222\_25637 -NCPU=6 -ICPU=3  
-NO\_XLPTX -NO\_MMAR -LARGESQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hc9:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sbs:\*  
12: gb\_ey:\*  
13: gb\_un:\*  
14: gb\_vt:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_cm:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_scs:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_dln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_ay:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	909.5	79.9	1022	8	PSMYB26	Y11105 P. sativum m
2	871	59.0	910	8	AY086615	AY086615 Arabidops
3	667	58.6	645	8	AF175987	AF175987 Arabidops
4	653.5	57.4	1033	8	AB005888	AB005888 Arabidops
5	585	51.4	1127	8	AY088761	AY088761 Arabidops
6	570	50.1	872	8	AB058642	AB058642 Lilium hy
7	520	45.7	1295	8	CPJ33917	U33917 Craterostig
8	519.5	45.7	105543	8	ATKCO09325	AC009325 Arabidops
9	518	45.5	1343	8	AY026332	AY026332 Oryza sat
10	515.5	45.3	1117	8	AF510112	AF510112 Craterost
11	514.5	45.2	13342	8	AP003140	AP003140 Oryza sat
12	514.5	44.8	145491	8	AP002883	AP002883 Oryza sat
13	510	45.2	1423	8	AF622733	AF622733 Arabidops
14	497	43.7	1140	8	AF334815	AF334815 Arabidops
15	485	42.6	931	8	AY008377	AY008377 Arabidops
16	479	42.1	853	8	AF474132	AF474132 Sorghum b
17	476.5	41.9	126599	2	AP003816	AP003816 Oryza sat
18	464	40.8	817	8	AY133705	AY133705 Arabidops
19	461.5	40.6	1033	8	AF034133	AF034133 Gossypium
20	458	40.2	107816	8	AC016447	AC016447 Arabidops
21	453.5	39.9	1070	8	ATU62743	U62743 Arabidops
22	446.5	39.2	158826	8	AP003607	AP003607 Oryza sat
23	446.5	39.2	197674	2	AP004367	AP004367 Oryza sat
24	440	38.7	1262	8	AF371980	AF371980 Arabidops
25	435.5	38.3	1558	8	CPJ33916	U33916 Craterostig
26	427.5	37.6	1024	8	AF062894	AF062894 Arabidops
27	427.5	37.6	114652	2	AC118288	AC118288 Oryza sat
28	426.5	37.5	134673	8	AP002860	AP002860 Oryza sat
29	425.5	37.4	2205	8	AB052240	AB052240 Arabidops
30	422	37.1	197135	2	AC124961	AC124961 Medicago
31	416.5	36.6	2131	8	PHMYBPH33	Z13998 P. hybri
32	415.5	36.5	771	8	AF272733	AF272733 Arabidops
33	413.5	36.3	1057	8	PHMYBPH22	Z13997 P. hybri
34	413.5	36.3	1155	8	AY059820	AY059820 Arabidops
35	411.5	36.2	90077	8	AP004915	AP004915 Lotus jap
36	410	36.0	960	8	AF175991	AF175991 Arabidops
37	410	36.0	1032	8	AY081512	AY081512 Arabidops
38	408	35.9	2220	8	AB052242	AB052242 Arabidops
39	408	35.9	2222	8	AB052245	AB052245 Arabidops
40	407	35.8	967	8	ATWYBRTF	Z68158 A. thaliana
41	406.5	35.7	1132	8	AY096523	AY096523 Arabidops
42	406.5	35.7	1512	8	AY063939	AY063939 Arabidops
43	405	35.6	2221	8	AB052244	AB052244 Arabidops
44	405	35.6	2221	8	AB052249	AB052249 Arabidops
45	404	35.5	795	8	AB029160	AB029160 Glycine m

RESULT 1

ALIGNMENTS

```

PSMYB26
LOCUS       PSMYB26                1022 bp    mRNA    linear    PLN 09-FEB-1998
DEFINITION  P.sativum mRNA for Myb-like protein (Myb26).
ACCESSION   Y11105
VERSION     Y11105.1
KEYWORDS    Myb-like protein; Myb26.
SOURCE      Pisum sativum.
ORGANISM    Pisum sativum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae;
Pisum.
REFERENCE   1 (bases 1 to 1022)
AUTHORS     Ullmar, A. and Strommer, J.
TITLE       Myb26: a MYB-like protein of pea flowers with affinity for
JOURNAL     promoters of phenylpropanoid genes
MEDLINE     Plant J. 12 (6), 1273-1284 (1997)
PUBMED      9450341
REFERENCE   2 (bases 1 to 1022)
AUTHORS     Strommer, J.N.
TITLE       Direct Submission
JOURNAL     Submitted (05-FEB-1997) J.N. Strommer, University Of Guelph,
Molecular Biology And Genetics, Guelph, Ontario N1G 2W1, CANADA
FEATURES
source
1..1022
/organism="Pisum sativum"
/db_xref="taxon:3888"
/feature_type="flower bud"
92..745
/codon_start=1
/product="Myb26"
/protein_id="CAA71992.1"
/db_xref="GI:1841475"
/db_xref="SPTRMBL:P93474"
/translation="MDKKPCNSODEPVRKGPWTMEIDILINVTIANHGGVGNLAK
AAGLKRKGSCLRLMNLVLRPVRGNITPEEDLIMELHSGKGRWSTLAKLIGRT
DNEIKNFMFRIOKHIKQVNDPNCOPKQMSLEINDHHHHHPHDSQSVNLVPEM
ETSPSTGYCTLEPPTGPTFTINDHHSNCCANDNNNTWMSMEDTMSWOLNGD"
BASE COUNT  370 a 170 c 161 g 321 t
ORIGIN
Alignment Scores:
Pred. No.:      1,73e-85      Length:      1022
Score:          909.50        Matches:     174
Percent Similarity: 83.18%    Conservative: 9
Best Local Similarity: 79.09%  Mismatches: 20
Query Match:    79.92%       Indels:      17
DB:             8            Gaps:        5
US-10-021-811-36 (1-206) x PSMYB26 (1-1022)
QY 1 MetAspLysIysGInGInCysLysThrSerGlnAspProGluValArgLysGlyProTTP 20
Db 92 ATGACAAAAA---CCCTGCAACTATCTCAAGATCTCGAAGTGAAGAAAGGCGCATGG 148
QY 21 ThrMetGlnGluAspLeuIleuMetAsnTyrIleAlaAsnHisGlyGluGlyValTTP 40
Db 149 ACCATGGAAGAAAGCTGATTTGATCAATTAATTAATTCATGTCGAAGGTGTTTGG 208
QY 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTTP 60
Db 209 AATTCTTACCAAGAGCTGCTGCTTAAACGACGGAAGAGTTCGACGGCTTCGATGG 268
QY 61 LysAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
Db 269 TTTAACTATCTTCTGCTCCAGATGTTAAGAGGAAATATTAACCTGAGGAACAACCTTTGG 328
QY 81 IleMetGlnLeuHisAlaLysTyrGlyAsnArgTTPSerLysIleAlaLysHisLeuPro 100
Db 329 ATCATGGAACCTTATTCATCAAGTGGGAAATAGGTGCTCCAAATATTCGAAAGCATCTTCA 388
QY 101 G-ArgThrAspAsnGlnIleLysAsnTyrTTPArgThrArgIleGlnLysHisIleLys 120

```

```

Db 389 GGAAGAACGACAAATGATTAAGAATTTTGGAGACTAGATCAAAAGACATTAAG 448
QY 121 GlnAla-----GluAsnPhengInGInSerSerAsnAsnSerGluIle 135
Db 449 CAAGTTGATTAACCTTAATCAACAATAATTTTCGCAAAAATATGAGT-----TTAGAGATA 502
QY 136 AsnAsp-----HisGlnAlaSerThrSerHisValSerThrMetAla 149
Db 503 AATGATCATCACCATCATCATCTCAACCAAGATGAGCAAGTTTCAATTGGTA 562
QY 150 GluProMetGluMetLysSerProProCysTyrGlnGlyMetLeuGluProPhseThr 169
Db 563 GAGCCATGGAACCTTATCTCCCACTTCATATCAAGAACTTGGACCATTTCCAACT 622
QY 170 GlnPheProThrIleAsnProAsp-----GlnSerCysCysThrAsnAspAsn 186
Db 623 CAATTTCAACTATTAATTAATATATCATCATCAAAACTCAAAATGTTGTGGCAATGACAAAC 682
QY 187 AsnAsnIleAsnTyrTTPSerMetGluAspSerTTPSerMetGlnLeuAsnGlyAsp 206
Db 683 AATTAACAACAATTAATTTGAGAGATGAGAGATATCTGGTCAATGCAATTACTCAATGGAGAT 742
RESULT 2
LOCUS       AY086615                910 bp    mRNA    linear    PLN 25-JUN-2002
DEFINITION  Arabidopsis thaliana clone 262460 mRNA, complete sequence.
ACCESSION   AY086615
VERSION     AY086615.1
KEYWORDS    FLI CDNA.
SOURCE      Thale cress.
ORGANISM    Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE   1 (bases 1 to 910)
AUTHORS     Haas, B.J., Volkovskiy, N., Town, C.D., Troukhan, M., Alexandrov, N.,
Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.
TITLE       Full-length messenger RNA sequences greatly improve genome
JOURNAL     annotation
REFERENCE   2 (bases 1 to 910)
AUTHORS     Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
Feldmann, K.
TITLE       Direct Submission
JOURNAL     Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
Malibu, CA 90265, USA
COMMENT     This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
sequence is considered to be 3'-truncated if it lacks the
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the WS or Laer ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
Geneset carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.
FEATURES
source
1..910
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="262460"

```

CDs	110.. 754	/codon_start=1	/product="putative transcription factor MYB24"
		/protein_id="A063674.1"	
		/db_xref="GI:21554734"	
		/translaton="MEKRSSSGSGSDAEVRKGPWTMEEDLILINVAIHGEGVNS	
		LASAGLAKTGKSCRLRWLNYLRPDVYRGNTPEEQLTIETLAKGNWSKLAKLP	
		GRNDIEKNFWRTKIOKYLIIKSGEITTVGSSEFIIHHAITSHVANDQETWDVYSP	
		TTSYQAHASNINQOALNENYIPBESSITMPLSVQSEQNTWSVDDLPMNIYNQN"	
BASE COUNT	307 a	169 c	185 g 249 t
ORIGIN			
Alignment Scores:			
Pred. No.:	1,13e-60	Length:	910
Score:	671.00	Matches:	133
Percent Similarity:	72.90%	Conservative:	23
Best Local Similarity:	62.15%	Mismatches:	50
Query Match:	58.96%	Indels:	8
	8	Gaps:	4
US-10-021-811-36 (1-206) x AY066615 (1-910)			
Qy	1	MetApLySvLysGlnGlnCys-----LysThrSerGlnAspProGlnValArgLysGly	18
Db	110	ATGGAGAAAGAGAAAGTAGTGTTCTGATCAGAGATGCAGAGTAGAAGAAAGG	169
Qy	19	ProTPRThrMetGluGluAspLeuIleLeuMetAenTyriIeAlaAsnHISGlyGluGly	38
Db	170	CCATGACGATGAGAGAAAGATTGATTGTTCTCATCAATTATGCGCAATCATGTGAAAGT	229
Qy	39	ValTPaenSerLeuAlaLysAlaIaGlyLeuLysValGsnGlyLysSerCysArgLeu	58
Db	230	GTTTGGAACTCTCTCGCAAAATCTGCAGAGACTAAACCACCGGAGAAAGTTGCGGCTC	289
Qy	59	ArgTPLeuAsnTyrrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGln	78
Db	290	CGGTGCGTAAGTACTCTCGACCTGATGCGACGGAGAAATATACACCGAGAGAACG	349
Qy	79	LeuLeuIleMetGluLeuHISAlaLysTyrrGlyAsnArgTyrrSerLysIleAlaLysHis	98
Db	350	CTCACCATCATGAACTTCATGCAAAATGGGAAATAGGTGTCAAAATTCGCAAGCAT	409
Qy	99	LeuProGlyArgThrAspAsnGluIleLysAsnTyrrTPArgThrArgIleGlnLysHis	118
Db	410	TTACACGAGAGAGACGACATAGATATAAGAACTTTTGAGAGACTTAAGATCCAGAAATAC	469
Qy	119	IleLysGlnAlaGluAsnPhelGlnGlnIleSerSerAsnAsnSerGluIleAsnAspHis	138
Db	470	ATCATCAAGACGCGAGAAACAGACACCGTTGGATCACAAGCTCCGAGTTTATAACAT	529
Qy	139	GlnAlaSerThrSerHisVal---SerThrMetAlaGluProMetGluMetTyrrSerPro	157
Db	530	CATGCGACACACGACCATCATGATGATGATCTCAAGAAACCATGATATGATTCTCCA	589
Qy	158	ProCysTyrrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsn-----	175
Db	590	ACGACGCTCGTATCAACATCCACGAATATTATATACGACACTTAATATTATGTAATTATG	649
Qy	176	ProAspGlnSerSerCys-----CysThrAsnAspAsnAsnIleAsnTyrrTP	192
Db	650	CCGTAAATCCAGTTCCGATCATGATGATGCCATTATCTGTGATCATCAATCGAACAATATTGG	709
Qy	193	SerMetGluAspSerTyrrSerMetGlnLeuLeuAsnGlyAsp	206
Db	710	AGGTGATGATGATCTTTGGCCCATGAATATATATATGATGAT	751
RESULT 3			
AF175987	645 bp	mRNA	linear
LOCUS	AF175987	complete cds.	PLN 30-AUG-2001
DEFINITION	Arabidopsis thaliana putative transcription factor (MYB24) mRNA,		
ACCESSION	AF175987		
VERSION	AF175987.1	GI:5823306	

KEYWORDS	Arabidopsis thaliana.
SOURCE	Arabidopsis thaliana
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE	1 (bases 1 to 645)
AUTHORS	Kranz, H.D., Denekamp, M., Greco, R., Jin, H., Leyva, A., Meisner, R.C., Petroni, K., Urraquin, A., Bayan, M., Martin, C., Smeekens, S., Tonelli, C., Paz-Ares, J. and Weisshaar, B.
TITLE	Towards functional characterisation of the members of the R2R3-MYB gene family from Arabidopsis thaliana
JOURNAL	Plant J. 16 (2), 263-276 (1998)
MEDLINE	99056848
PUBMED	9839469
REFERENCE	2 (bases 1 to 645)
AUTHORS	Stracke, R., Werber, M. and Weisshaar, B.
TITLE	The R2R3-MYB gene family in Arabidopsis thaliana
JOURNAL	Curr. Opin. Plant Biol. 4 (5), 447-456 (2001)
MEDLINE	21481677
PUBMED	11597504
REFERENCE	3 (bases 1 to 645)
AUTHORS	Stracke, R. and Weisshaar, B.
TITLE	Direct Submission
JOURNAL	Submitted (06-AUG-1999) Dept. Biochemie, Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln D-50829, Germany
FEATURES	Location/Qualifiers
source	1..645
	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"
	/note="ecotype: Columbia"
gene	1..645
	/gene="MYB24"
	/note="R2R3-MYB family member"
CDS	1..645
	/codon_start=1
	/product="putative transcription factor"
	/protein_id="AA053092.1"
	/db_xref="GI:5823307"
	/translation="MEKRSSGSGSGSDAYRKQPTMEEDLIINIANHREGVNSI/LAKSAGLKRTGSKCRKLWNLARPDVRGNIITPEBQFLIEMHAKGNRKSIAKHPGRDNEIKFWRTKIQRYIKSGPTTVNGQSSEFIINHAATTSVMNDTDETMYSPTTSYQASNIINQOLANTYNGYVESGSIIMPLSVDSQEQYMSVVDLWPMNTYNGN"
BASE COUNT	219 a 125 c 153 g 148 t
ORIGIN	
Alignment Scores:	
Pred. No.:	1.95e-60
Score:	667.00
Percent Similarity:	72.43%
Best Local Similarity:	61.68%
Query Match:	58.61%
DB:	8
	Gaps: 4
US-10-021-811-36 (1-206) x AF175987 (1-645)	
QY	1 MetaaPlybLysGInGInCys-----LysThrSerGInaSPProGUValaArgHysGly 18
Db	1 ATGGAGAAAAGAAAGTAAGTGGTGGTCTGGATACAGAGATGCAGAGGTGAGAAAAGGG 60
QY	19 ProTPTThMeGlGInaSPLeuLeileuMeCaenTYrIleAlaAsnHieGlyGluGly 38
Db	61 CCAATGACGATGGAAGAAAGATTGATTCTCATCATATATATGCCAATCATGTGAAGGT 120
QY	39 ValTPaasSerLeuAlaLysAlaAaGlyLeuLysaRgaemGlyLysSerCyARglau 58
Db	121 GTTTGGAACTCTCGCCAAATCTCAGAGACATAAACGACCGGGGAAAAGTTGCCGGCTC 180
QY	59 ArgTPTLeuAsnTYLeuLysProAspValaRGAyGlyAAsnIleThProGUluGlu 78
Db	181 CGATGGCTGAACCTCTCGACCTTAATGGCGACGGGGAATATATACACCAAGAGACAG 240

QY 79 leuLeuIleMetGluLeuHisAlaIleTyrGlnIleAsnArgTyrSerIleAlaIleHis 98  
 Db 241 CTCACCATCATGGAACCTTCATCAAAATGCGCAATAGGTGTCAAAAATGCAAGCAT 300  
 QY 99 leuProGlnIleTyrThrAspAsnGlnIleIleAsnTyrTTPAqThrArgIleGlnIleHis 118  
 Db 302 TTACCAAGGAAGAACCAATGATTAAGAAATTTTGGAGCACTAAGATCCAGAAATAC 360  
 QY 119 leuIleGlnAlaGluAsnProGlnIleGlnIleSerSerAsnAsnSerGlnIleAsnAspHis 138  
 Db 361 ATCATCAAGAGCGGAGAAACGACGCGTGGATACAAAGCTCCGATTTATAACCAT 420  
 QY 139 GlnIleSerThrSerHisIleVal---SerThrMetAlaGluProMetGluMetTyrSerPro 157  
 Db 421 CATGGACACACGACCATGATCATGATATCTCAAGAAACCATGATATGATATCTCCA 480  
 QY 158 ProCysTyrGlnIleMetLeuGlnIleProPheSerThrGlnIlePheProThrIleAsn----- 175  
 Db 481 AGCAGGTGTTCAACATACGACCAATATTAATCAGCGCTTAATTATGTTAATTATGTG 540  
 QY 176 ProAspGlnIleSerCys-----CysThrAsnAspAsnAsnAsnIleAsnTyrTyr 192  
 Db 541 CCGTATCCGGTTCGATCATGATGATGATTCGTTGATCATATCCGAACAAACTATTGG 600  
 QY 193 SerMetGluAspSerTyrSerMetGlnIleuLeuAsnGlyasp 206  
 Db 601 AGCGTCATGATCTTGGCCCATGATATATATATATGTTAAT 642

RESULT 4  
 AB005888 1033 bp mRNA linear PLN 05-FEB-1999  
 LOCUS Arabidopsis thaliana mRNA for ATMVB3, complete cds.  
 DEFINITION AB005888  
 VERSION AB005888.1 GI:2280527  
 KEYWORDS ATMVB3.  
 SOURCE Arabidopsis thaliana (strain:Columbia) cDNA to mRNA.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (sites)

REFERENCE  
 AUTHORS Noji,M., Urago,T., Shinozaki,K.Y. and Shinozaki,K.  
 TITLE Molecular cloning of two cDNAs encoding novel myb homologs from  
 Arabidopsis (Accession Nos. AB005888 and AB005889) (PGR98-111)  
 JOURNAL Plant Physiol. 117, 720 (1998)  
 REFERENCE  
 AUTHORS Noji,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-JUL-1997) Masaaki Noji, Chiba University, Faculty of  
 Pharmaceutical Sciences; Yayoi-cho 1-33, Inage-ku, Chiba,  
 263, Japan (E-mail:mmoji@chiba-u.ac.jp, Tel:+81-43-290-2906,  
 Fax:+81-43-290-2905)

FEATURES  
 source location/Qualifiers

1..1033  
 /organism="Arabidopsis thaliana"  
 /strain="Columbia"  
 /db\_xref="taxon:3702"  
 1..1033  
 /gene="Atmyb3"  
 115..795  
 /gene="Atmyb3"  
 /codon\_start=1  
 /product="ATMYB3"  
 /protein\_id="BA021618.1"  
 /db\_xref="GI:2280528"  
 /translation="MEKRGSGSSGSSGSAEAEVRKGPWTMEEDLILINYLANHGQV  
 VNSIAAGIKRTGKSCRLRWNLNLRPDVRANITPERDLIMEHAHAKRWKSLAK  
 HLEPRTDNEIKNCRTRIKYIKOSGVTTTSSVGSHSSEINDQAASSSHAVFCTQD  
 QAMETYSPTPTSYGHTHMEFNKGNISAAAVTATVDVPPMTVDQGTGNYWGMDDIWS  
 SHALNGN"

BASE COUNT 362 a 182 c 208 g 281 t

Alignment Scores:  
 Pred. No.: 8 766-59 Length: 1033  
 Score: 653.50 Matches: 137  
 Percent Similarity: 71.83% Conservative: 16  
 Best Local Similarity: 64.32% Mismatches: 45  
 Query Match: 57.43% Indels: 15  
 DB: 8 Gaps: 6

US-10-021-811-36 (1-206) x AB005888 (1-1033)

QY 9 ThrSerGlnAspProGlnValAlaIleGlyGlnIleProTyrThrMetGlnGluAsnIleLeu 28  
 Db 154 TCATCGACAGAACGACAAAGTGAAGAAAGGACCATGACATGAGAAAGATCTTATCTT 213  
 QY 29 MetAsnTyrIleAlaAsnHisGlnIleGlnIleValTTPAsnSerLeuAlaIleGly 48  
 Db 214 ATCACTATATGCGCAACCAACGCGCATGGTGTGGAAATTCCTCGCAAACTGACAGT 273  
 QY 49 LeuIleArgAsnGlnIleIleSerCysArgLeuArgTyrPLeuAsnTyrIleuArgProAspVal 68  
 Db 274 CTAAACGAAACCGGGAAGTTGCGCGCTCCGGTGGTGAATCTCTCGCCCGACGTA 333  
 QY 69 ArgArgGlnAsnIleThrProGlnIleGlnIleuLeuIleMetGluLeuHisAlaIleTyr 88  
 Db 334 CAGCGGGAAACATCATCTCCAGAAAGCACTATCATCATGAGAACTTCATGCTTAAGTGG 393  
 QY 89 GlnAsnArgTyrSerIleIleAlaIleHisIleuProGlnIleArgThrAspAsnGlnIleIle 108  
 Db 394 GGAAACAGGTGTCGAAATATGCGCAACATCTTCCAGGAAGACGACGAAAGATCAA 453  
 QY 109 AsnTyrTTPAqThrArgIleGlnIleHisIleIleGlnAlaGlu-----AsnPhe 125  
 Db 454 AATTCTGTAGGACAAAGATTCAAAAATCATCAAGCATGATGATGTAACAACAACATCG 513  
 QY 126 GlnGlnIleSerSerAsnAsnSerGlnIleAsnAspHisGlnIleSerThrSerHis--- 144  
 Db 514 TCCGTTGATCTCATCATACGATCGAGATCAACATCAAGTCAGACGACGTCAGCCAT 573  
 QY 145 -----ValSerThrMetAlaGluProMetGluMetTyrSer-----ProProCysTyr 160  
 Db 574 AATGCTTTTGTGCACAAAGATCAACGATGAGACTTATCTCTACACGACATCATAT 633  
 QY 161 GlnIleMetLeuGlnIleProPhe-----SerThrGlnIlePheProThrIleAsn 175  
 Db 634 CAACATACCATATGATGATTTCAACTATGTAATCTATTCGCGCGGACGACGACCAAC 693  
 QY 176 ProAspGlnIleSerCysCysThr---AsnAspAsnAsnAsnIleAsnTyrTyrSerMet 194  
 Db 694 GTGGATTATCCAGTACCGATGACCGTTGATGATCAAAACCGGTGAAAACTATTTGGGCATG 753  
 QY 195 GlnAspSerTyr---SerMetGlnIleuLeuAsnGlyasp 206  
 Db 754 CATGATATTTGGTCAATCATGATCATTTATGATGTTAAT 792

RESULT 5  
 AY088761 1127 bp mRNA linear PLN 27-JUN-2002  
 LOCUS Arabidopsis thaliana clone 94595 mRNA, complete sequence.  
 DEFINITION AY088761  
 ACCESSION AY088761  
 VERSION AY088761.1 GI:21407535  
 KEYWORDS F11 cDNA.  
 SOURCE F11 cDNA.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 1127)

REFERENCE  
 AUTHORS Haas,B.U., Volkovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,  
 Feldmann,K.A., Flavel,R.B., White,O. and Salzberg,S.L.  
 TITLE Full-length messenger RNA sequences greatly improve genome  
 annotation  
 JOURNAL Genome Biol. (2002) In press  
 REFERENCE 2 (bases 1 to 1127)

AUTHORS Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.  
TITLE Full-length cDNA from Arabidopsis thaliana  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 1127)  
AUTHORS Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.  
TITLE Direct Submission  
JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA  
COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or later ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genes carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES

source

1..1127  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/clone="94595"  
274..894  
/codon\_start=1  
/product="putative transcription factor"  
/protein\_id="AA067076.1"  
/db\_xref="GI:21618026"  
/translation="METITMKKGRVATTSQKEESTVRKGPWTEEDFILFNLYLHSGRLMNSVAKASGLKRTGKSCRLRWLNYPVPRRNITTEBQLIIOHAAQGNRW SKIAHLPGKRDNEIKNMPRTKIQRMKVSSEMMNQHSCNSGSSGTTTGGSSGK AIDTAEFSQAKTTTFNVVEQSNENYMNVMGIDLPVHLANGDHV1"

BASE COUNT 353 a 203 c 241 g 330 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1,35e-51 Length: 1127  
Score: 585.00 Matches: 121  
Percent Similarity: 65.14% Conservative: 21  
Best Local Similarity: 55.50% Mismatches: 44  
Query Match: 51.41% Indels: 32  
DB: Gaps: 6

US-10-021-811-36 (1-206) x AY088761 (1-1127)

OY 1 MetAaplyblygInGlnCyls-----ThrSerGlnaPProGlu----- 14  
DB 286 ATGAAGAAGAAAGAGAGTGAAGCAATTAACGTACAGAAAGAAAGAAAGAAC 345  
OY 15 ValAarglyblyPProTThrMetGlnuAspLeuileuMetAsnTyrlleAlaasn 34  
DB 346 GTGAGAAAGAGAGCTTGAAGCTTGAAGAGAGTTTCATCTCTTTAATTAATCTTAA 405  
OY 35 HisgIygluGlyValTTrpAsnSerLeuAlaYsAlaagIyLeuLyAsnGlyYls 54  
DB 406 CATGGTGAAGAGCTTTTGAAGCTCTGCGCAAAAGCTCGGTCTAAACGTAACGAAAA 465  
OY 55 SerCySaGleuAargTTrpLeuAsnTyrlleuAargProAspValaArglyAsnlethr 74  
DB 466 AGTTTGCGGCTCGGTGCGTGAACATATCTCCACAGATGTCGCGGAGGAAACCTAAC 525  
OY 75 ProGluGlnGlnLeuLeuileuMetGlnuLeuHisAlaYsTTrpGlyAsnAargTrpSerly 94  
DB 526 ACGAAGAAACAGCTTTTATCATCTACGCTTCACTAAGCTTGAAGAAACAGGTGCGAAG 585

OY 95 IleAlaYsHisLeuProGlyAargThraspaengulileYsAsnTyrlTrpAargThra 114  
DB 586 ATTCGGAAGCATCTTCCGGAGAACGACAGATTAAGAACTTCTTGAGAGCAAAAG 645  
OY 115 IleGlnYsHisIleYsGlnAlaGlu-----AanPheGlnGlnInsSerSer 130  
DB 646 ATTCAGAGACATGAAAGTGCATCGGAAAAATATGATGAATCATCAATCATTTGTC 705  
OY 131 AsnAsnSerGluIleAsn-----AspHisGlnAlaSerThrSerHisValSerThrMet 148  
DB 706 GGAACATCAACAGAGCTCGGAGATGACGACGACGACAGGACCTCCGCAAGACATAGACAG 765  
OY 149 AlaGluProMetGlnuMetYsSerProProCysTyrlGlnGlyMetLeuGlnuProPheSer 168  
DB 766 GCTGAGAGCTTCTCT-----CAGCGAAAGCG 792  
OY 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsn 188  
DB 793 ACAGCTTTAATGTGCTG-----GAACACAGTCAAC 825  
OY 189 IleAsnTyrlTrpSerMetGlnuAspSerTrpSerMetGlnuLeuAsnGlyAsp 206  
DB 826 GAGAAATTACTGGAAGCTTGAGATCTGTGCGCCGTCACCTTCTTAATGATGAC 879

RESULT 6  
AB058642 872 bp mRNA linear PLN 04-APR-2001  
LOCUS Lilium hybrid division I Lmhyb mRNA, complete cds.  
DEFINITION AB058642  
ACCESSION AB058642.1 GI:13537529  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Lilium hybrid division I (cultivar:Montreux) tepal cDNA to mRNA.  
Lilium hybrid division I  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;  
Lilium; Lilium hybrid cultivars.  
1 (sites)  
REFERENCE  
AUTHORS Nakatsuka, A., Izumi, Y. and Yamagishi, M.  
TITLE Isolation and characterization of the genes related to anthocyanin biosynthesis in Asiatic hybrid lily  
JOURNAL Unpublished  
2 (bases 1 to 872)

AUTHORS Nakatsuka, A., Izumi, Y. and Yamagishi, M.  
TITLE Direct Submission  
JOURNAL Submitted (24-MAR-2001) Akira Nakatsuka, Shimane University, Faculty of Life and Environmental Sciences, Matsikawatsu 1060, Matsue, Shimane 690-8504, Japan  
(E-mail:nakatsuka@life.shimane-u.ac.jp, Tel:81-852-32-6502)

FEATURES  
source  
1..872  
/organism="Lilium hybrid division I"  
/cultivar="Montreux"  
/db\_xref="taxon:156532"  
/tissue\_type="tepals"  
1..872  
/gene="Lmhyb"  
80..634  
/gene="Lmhyb"  
/codon\_start=1  
/protein\_id="BAB40790.1"  
/db\_xref="GI:13537530"

gene  
CDS  
80..634  
/gene="Lmhyb"  
/codon\_start=1  
/protein\_id="BAB40790.1"  
/db\_xref="GI:13537530"

BASE COUNT 269 a 177 c 213 g 213 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.65e-50 Length: 872  
Score: 570.00 Matches: 118  
Percent Similarity: 64.29% Conservative: 17  
Best Local Similarity: 56.19% Mismatches: 39

Query Match: 50.09% Indels: 36  
 DB: 8 Gaps: 5  
 US-10-021-811-36 (1-206) x AB058642 (1-872)

```

QY 1 MetApblyblygInglncylserThrserrInaPProgluValArgyleGlyProTyr 20
DB 80 ATGGACAGGAA-----GTAGGAGCTATGAAAGATGTTGAGGAGGAAAGACCATGG 133
QY 21 ThrMetGluAspLeuIleuMetAsnYrlllealaaenHsGlyGluGlyValTyr 40
DB 134 ACATGAGAAAGAGATCTCATCTCTCATCATGACATGACCAACATGGGACGAGATCTGG 193
QY 41 InsSerLeuAlaLysAlaLysIleuLysArgAnGlyLysSerCysArgLeuArgTyr 60
DB 194 AACAGCTGGCTCGCATCGCCCGCATGGAAGAGACTGGAAGAAGCTGACAGCTTCGTTGG 253
QY 61 LeuAsnYrleuArgProAspValArgArgGlyAsnIleThrProGluGluGluLeu 80
DB 254 CTGAATACCTGAGGCTGATGTTGTCGGGGTAACTACACACAGAGACCATCTTCG 313
QY 81 MetGluLeuHsAlaLysIleuLysArgTyrSerLysIleAlaLysHsLeuPro 100
DB 314 ATCATGAACTTCACGACAGATGGGTAACAGATGTCGAAATTCGAAAGAGACTCCA 373
QY 101 GlyArgThrAspAsnGluIleLysAsnYrTrpArgThrArgIleGluLysHsIleLys 120
DB 374 GCAGCAACAGAAATCAAGAACTTCGAGAGACCAAGTCCAGAAAG-----424
QY 121 GluAlaGluAsnPhgInglncInserSerAsnAsnSerGluIleAsnAspHsGlnAla 140
DB 425 -----AAATCCAAAGATGGCGAATCATCTGACGGGCGCAATC 460
QY 141 SerThrSerHsValSerThrMetAlaGluProMetGlu-----153
DB 461 TCCATCTTCCTGGATGAACCTAGTCCATGCAAGATGCAAGATGTCGATGTTGGTGTG 520
QY 154 -----MetYrSerProProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPhe 171
DB 521 CAGACGACCTACGATCAAGCCGACGCTTCGAAACAACCTTGACGATTTGAAATCTCTTC 580
QY 172 ProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsnIleAsnTyr 191
DB 581 -----GAGCTGGCGAT-----AACTCC 598
QY 192 TrpSerMetGluAspSerTrpSerMetGln 201
DB 599 CTGTCCGTCGAGAACTGGGCTATGCAG 628

```

RESULT 7  
 CPU33917  
 LOCUS 1295 bp mRNA linear PLN 17-APR-2001  
 DEFINITION Craterostigma plantagineum myb-related transcription factor (cpm7)  
 mRNA, complete cds.  
 U33917  
 ACCESSION U33917.1 GI:1002799  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Craterostigma plantagineum.  
 Craterostigma plantagineum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Lamiales; Lamiales incertae sedis;  
 Toreniaeae; Craterostigma.  
 1 (bases 1 to 1295)  
 Iturriaga, G., Leyna, L., Villegas, A., Gharaibeh, R., Salamini, F. and  
 Bartels, D.  
 A family of novel myb-related genes from the resurrection plant  
 Craterostigma plantagineum are specifically expressed in callus and  
 roots in response to ABA or desiccation  
 Journal Mol. Biol. 32 (4), 707-716 (1996)  
 MEDLINE 97134962  
 PUBMED 8980522  
 REFERENCE 2 (bases 1 to 1295)

AUTHORS Iturriaga, G.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-AUG-1995) Gabriel Iturriaga, Plant Molecular Biology,  
 Instituto de Biociencia, Av. Universidad #2001, Cuernavaca, MOR  
 62210, Mexico

FEATURES  
 source  
 1..1295  
 /organism="Craterostigma plantagineum"  
 /db\_xref="taxon:4153"  
 gene  
 1..1295  
 /gene="cpm7"  
 125..1132  
 /gene="cpm7"  
 /note="putative DNA-binding protein; Description: myb-like  
 gene; myb-related transcription factor"  
 /codon\_start=1  
 /product="cpm7"  
 /protein\_id="AAB58314.1"  
 /db\_xref="GI:1002800"  
 /translation="MNQOQVYKSKRNQINSEDDDDSSDLRGPWVDEDFLLINYIA  
 HHGGRNMLARFAGLKRKTSGLRWLNTYLRPVRRGNLTLEOLLILEHSRNGR  
 WSKIAOHLGRDNEIRNWRTRVOKAKOLKQVNSKOPEDMRVYIMPRVLROR  
 AATDDGAPPLASSSPSSAMNTACYSTMAADHROPLMAYYAMTHSTMLTA  
 QENSSTVASSPSSISSETEAYNTANHOVINGADHQIDSTTISYGMNDPAINT  
 VGNNGNSDQLQMGSVDDRRSNEQMMMDVDVNGSSSDKNNLLMNVDVWFLQDFSSC  
 F"

BASE COUNT 405 a 286 c 279 g 325 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 9.6e-45 Length: 1295  
 Score: 520.00 Matches: 104  
 Percent Similarity: 62.38% Conservative: 27  
 Best Local Similarity: 50.00% Mismatches: 31  
 Query Match: 45.69% Indels: 46  
 DB: 8 Gaps: 5

US-10-021-811-36 (1-206) x CPU33917 (1-1295)

```

QY 1 MetApblyblygInglncylserThrserrInaPProgluVal-----15
DB 125 ATGAACCAACAGCAGATGTTGTTAAAGTTTCCAAAGAACAAATTAACATGAAAGACAT 184
QY 16 -----ArgyleGlyProTyrTrpMetGluAspLeuIleuMet 29
DB 185 GATGATTCCTGGACCTCAAGAAAGACCGCTGATGATGAAAGACTTCACACTATC 244
QY 30 AsnYrlllealaaenHsGlyGluGlyValTrpAsnSerLeuAlaLysAlaLysLeu 49
DB 245 AACTACATCGCTCACCATGCGAAGGAGAGATGGAACCTCTTGGACGCTTTCGGGCTGG 304
QY 50 LysArgAsnGlyLysSerCysArgLeuArgTyrLeuAsnYrleuArgProAspValArg 69
DB 305 AACGAACTCGAAAGAGCTGACGATGATGATGTTAAACTTCTTGACACCGATCTTCGT 364
QY 70 ArgGlyAsnIleThrProGluGluGluLeuIleuMetGluLeuHsAlaLysIleu 89
DB 365 CAGAGAAATATCACACTCGAAGAAACAGCTTCTCATTTCTGAACTCCATTCCTGGGGGC 424
QY 90 AsnArgTyrSerLysIleAlaLysHsLeuProGlyArgThrAspAsnGluIleLysAsn 109
DB 425 AATAGTGTGTCGAAAGATGGCCCAACATTTGCTCGAAGACAGACATGATTAAGAAC 484
QY 110 TyrTrpArgThrArgIleGluLysHsIleLysGlnAlaGluAsnPhgInglncInser 129
DB 485 TACTGAGAAAGAGAGGTCCTCAAAAACATGCCAAACAGCTA-----AAGTGCAGAC 532
QY 130 SerAsnAsnSerGluIleAsnAsp-----137
DB 533 GTCAACACGAAACGTTAAAGACCAATGATGATCCTTGGATGCCAAGATTGGTGCAG 592
QY 138 -----HisGlnAlaSerThrSerHis-----ValSerThrMetAla 149

```



Db 593 AGATCCACCTGCGCCACCAACGACGCGCCACCCCTCTGCGCTCCGCTCC 652  
Oy 150 GUBPROMETGLUMETLYRSEPRPROCYETTYR----- 160  
Db 653 TCGCGGTCATCCGACATGATACAGCCTGCTACAGCACTGCCATGCTGCAGGCGATCAC 712  
Oy 161 -----GinglyMettLeuGluPro 166  
Db 713 CGTGTACAGTTATGCTATGTCGA 736

RESULT 8  
ATAC009325/c 105543 bp DNA linear PLN 24-JAN-2001  
LOCUS Arabidopsis thaliana chromosome III BAC F4P13 genomic sequence,  
DEFINITION complete sequence.  
ACCESSION AC009325  
VERSION AC009325.8 GI:12408717  
KEYWORDS HTG.  
SOURCE Arabidopsis thaliana.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 105543)  
AUTHORS Lin,X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B.,  
Rommig,C.M., Koo,H., Fujii,C.Y., Uteback,T.R., Barnstead,M.E.,  
Bowman,C.L., White,O., Nielsen,W.C. and Fraser,C.M.  
Arabidopsis thaliana chromosome III BAC F4P13 genomic sequence  
Unpublished  
2 (bases 1 to 105543)  
REFERENCE Lin,X. and Kaul,S.  
AUTHORS Direct Submission  
JOURNAL Submitted (16-ANG-1999) The Institute for Genomic Research, 9712  
Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org  
3 (bases 1 to 105543)  
REFERENCE Lin,X.  
AUTHORS Direct Submission  
JOURNAL Submitted (24-JAN-2001) The Institute for Genomic Research, 9712  
Medical Center Dr., Rockville, MD 20850, USA  
On Jan 24, 2001 this sequence version replaced gi:12280821.  
COMMENT Address all correspondence to:  
Xiaoying Lin  
The Institute for Genomic Research  
9712 Medical Center Dr.  
Rockville, MD 20850, USA  
e-mail: xlin@tigr.org  
BAC clone F4P13 is from Arabidopsis chromosome III and is near the  
molecular marker ILRI.  
The orientation of the sequence is from SP6 to T7 end of the BAC  
clone.  
Genes were identified by a combination of three methods: Gene  
prediction programs including GRAFT (available by anonymous ftp  
from artthur.epm.ornl.gov), GeneFinder (Phil Green, University of  
Washington), Genscan (Chris Burge,  
http://www.csb.stanford.edu/~chris/burge/GENSCANW.html), and NetPlantGene  
(http://www.cbs.dtu.dk/nepegene/obsnetgene.html), searches of the  
complete sequence against a peptide database and the Arabidopsis  
EST database at TIGR (http://www.tigr.org/cdb/at/est.html).  
Annotated genes are named to indicate the level of evidence for  
their annotation. Genes with similarity to other proteins are named  
after the database hits. Genes without significant peptide  
similarity but with EST similarity are named as 'unknown' proteins.  
Genes without protein or EST similarity, that are predicted by more  
than two gene prediction programs over most of their length are  
annotated as 'hypothetical' proteins. Genes encoding tRNAs are  
predicted by tRNAscan-SE (Sean Eddy,  
http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are  
identified by RepeatMasker (Arian Smit,  
http://ftp.genome.washington.edu/BM/RepeatMasker.html). Regions of  
genomic sequence that are not annotated as genes but have predicted  
exons by GRAFT are annotated as misc features.

FEATURES  
source

1..105543

/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="III"  
/map="ILRI"  
/clone="F4P13"  
/gene="P4P13.1"  
/note="P4P13.1"  
/join(47..145,499..597,951..1079,1232..1387,1575..3626,  
3721..4029)  
/gene="F4P13.1"  
/note="F4P13.1"  
/protein\_id="AA01531.1"  
/codon\_start=1  
/protein\_id="GI:6016705"  
/translation="MVSRLPDLPTITDLRLAAGAYDGTAEFLDYELIMSSIRMYAD  
QPDVDIVATLSEKFKSLYEAEVPLVQKLDYRKLKELTSAEMKEIKDIVSVNKL  
KAPMDGMAQALAESYKLVRRKGRKXGELTPASMBELTHLADVMEKDYWEFSAR  
RIILKLCEBELSSLVHQLRQCAEAIEMOQKLRLSLEWKNAMROEFTLAK  
AKVEPSILKEVGEPHNSGFPADMGCDPQPGQDGYTRDDETSTAYLKNQKSP  
LETDPQRESHVNRGSKTSSPRTTSSPGREHLPIADTSPVTNLPKDTSETLKS  
VERNETHSFNSNAVELPTADNASQASQELQACQDLSATSNELQNTQSRITSEQ  
LLKQSRDLFLGTGDSGLYWGCCPDPNPILVDGSLSLQKPVQADLIGKVSFPL  
HVDHGRRLFPWTYETETETISLVQVLMHDDILEKDRRSIIMWKRLRGDVQEK  
KQAOMLSAPFATGLETAKAMSEKRYPCIKLEMETLKGKTKYVAERKLCRCCK  
LESILPSMHLICHTFASDDEFDEHESKCIPISLATEKGRKXGELTSAEMKEIK  
YLVKSSAKQVYARIISNYSHELDSGLIRYQREESIISPHFERIKCFKFTKCDNRLVSE  
IGLISNGIPTFLPSSSTHLNDSVLISAKSNKPPGSGDGYTPAGETTVNGKES  
NMSFPRSVTDSHGPLDKPGLGFGFSGKQKKSSGSLKSCVVPQALRVTGKAL  
PGRFLKTNLLDMVDALPEALRPESKSPNRRRAMRVKSSQSYELVQATIVEDM  
IKTEYLKEMWYMSLSAAAKISTLSLSVFIIFSDAALIVDKPITPSPNPIDETPII  
SLPDQKQPVSDSQERSRVRSGKKRKEPGS"  
4160..4300  
/note="exon predicted by xgrail, quality good"  
complement(4496..6383)  
/gene="F4P13.2"  
/note="identical to homeobox-leucine zipper protein HAT5  
(HD-ZIP protein 5) (HD-ZIP protein ATHB-1) GB:002283  
(Arabidopsis thaliana)"  
complement(join(4496..4989,5292..5531,5619..5758,  
5938..6383))  
/gene="F4P13.2"  
complement(join(4639..4989,5292..5531,5619..5758,  
5938..6025))  
/gene="F4P13.2"  
/codon\_start=1  
/product="homeobox-leucine zipper protein HAT5 (HD-ZIP  
protein 5) (HD-ZIP protein ATHB-1)"  
/protein\_id="AA01532.1"  
/db\_xref="GI:6016706"  
/translation="MESNRPFPDPASHSNMFILGNIPVVGSGARSMMNMEETSK  
RBPFSSEPDLYDDPYDQPEKGRKLTTEVHLBSFETBNKLTBERKTLQAKKL  
GIQPRVAVWQFNRRARWTKQLERDYDLKSTYQULSNYSIVMDNDKLSVYSL  
TEKLOQGTETANPEGOVPEPQOLDPVYINAAITETRLSSGVSATLDDAQQLL  
DSQDSYFSPVFIQDMSNASDHDNDRSCFADVFVPTSPSHDHESLSAFWGP"  
complement(6749..6849)  
/note="exon predicted by xgrail, quality marginal"  
6892..6984  
/note="exon predicted by xgrail, quality  
marginal shadowexon"  
complement(6934..6974)  
/note="exon predicted by xgrail, quality good"  
8088..8181  
/rpt\_family="TAAAAn"  
complement(8901..8985)  
/note="exon predicted by xgrail, quality  
marginal shadowexon"  
9027..9064  
/rpt\_family="GAAAAn"  
complement(9068..9130)





Query Match: 45.30% Indels: 25  
 DB: 8 Gaps: 5  
 US-10-021-811-36 (1-206) x AF510112 (1-1117)

Oy 7 (YSLVSTHrserGlnAsp---ProGluValArgGlyProTTPHrMetGluGluAsp 25  
 Db 130 (YSLVSTHrserGlnAsp---ProGluValArgGlyProTTPHrMetGluGluAsp 189  
 Oy 26 (LeuLeuMetAsnTyrTlleAlaAsnHisGlyGluGluValTTPAsnSerLeuAlaAs 45  
 Db 190 (TTCACACTCATCACTCATCGCTCCACCATGCGCAAGAGATGGAATCTCTTGGACGT 249  
 Oy 46 (AlaAlaGlyLeuValArgAsnGlySerGlySerGlySerGlySerGlySerGlySerGly 65  
 Db 250 (TTCGCTGGCTGGAACGAACTGGAAGAGCGAGATGGAATGGAATGGAATGGAATGGA 309  
 Oy 66 (ProAspValAlaGargGlyAsnTlleThProGluGluGluLeuLeuLeuMetGluLeuHis 85  
 Db 310 (CCGACGTTCCGTCGAGAAATATATCACTCGAAGAACAGCTTTCATTCTTGAATCCAT 369  
 Oy 86 (AlaLysTTPGlyAsnArgTTPSerTyrTlleAlaLysHisLeuProGlyValArgThAspAsn 105  
 Db 370 (TCACTTGGGCAATAGGTGGTGGCAAGATTCCTCAACATTTGCTGGAAGAGACGACAAAT 429  
 Oy 106 (GluLeuLysAsnTyrTTPArgTThArgTlleGluHisHisTlleLysGluAlaGluAsnPh 125  
 Db 430 (GAGATAAAGAACTACTGAGAACAAAGGTCACAAAACATGCCAAACAGCTA----- 480  
 Oy 126 (GlnGlnGlnSerSerAsnSerGluLeuLeuAsnHis----- 138  
 Db 481 (---AAGTCCGACGTCACACGAAACAGTTCAAAAGACACCATGAGATACCTTTGGATGCCA 537  
 Oy 139 (-----GlnAlaSerThr-----SerHisValSerThr 147  
 Db 538 (AGATTGGTCGAGAGATCCAAAGCTTCCGCGCACACCGACGACGAGCGCGCGCGTGT 597  
 Oy 148 (MetLlaGluProMetGluMetTyrSerProProGlySerTyrGlnGlyMetLeuGlu----- 165  
 Db 598 (GCCCTCCGCGCTGCTCCGCGCATGAAATACAGCTGCTACAGCGCTGCTGCGTGCAGGC 657  
 Oy 166 (ProPheSerThrGlnPheProThrTlleAsnProAspGlnSerSerCysCysThrAsnAsp 185  
 Db 658 (GATCACCGTCGCTGATTATGCTGATGCCAGATTAACGACGACGACGACGACGCCAC 717  
 Oy 186 (AsnAsnAsnTle 189  
 Db 718 (AATTAATTCGATG 729  
 RESULT 11  
 AP003140 133242 bp DNA linear PLN 21-MAR-2002  
 LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,  
 DEFINITION BAC clone:OSJNBa0025P13.  
 ACCESSION AP003140  
 VERSION AP003140.2 GI:14164491  
 KEYWORDS  
 SOURCE Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,  
 clone:OSJNBa0025P13.  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE  
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.  
 TITLE Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, BAC  
 JOURNAL clone:OSJNBa0025P13  
 REFERENCE Published in Database (2001)  
 AUTHORS 2 (bases 1 to 133242)  
 TITLE Sasaki, T., Matsumoto, T. and Yamamoto, K.  
 JOURNAL Direct Submission  
 Submitted (31-JAN-2001) Takuji Sasaki, National Institute of  
 Agrbiological Resources, Rice Genome Research Program, Kamondai

## COMMENT

2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail:tsasaki@abrr.affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/,  
 Tel:81-298-38-7441, Fax:81-298-38-7468)  
 On May 21, 2001 this sequence version replaced gi:12641874.  
 Genes were predicted from the integrated results of the following:  
 (GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor  
 (October 1998 version). The genomic sequence was searched against  
 NCBI NonRedundant Protein database, nr  
 (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at  
 RGP. Protein homologies of the coding regions were searched against  
 NCBI NonRedundant Protein database with BLAST2.0. ESTs represent  
 the identified cDNA sequences using BLAST2.0 with the  
 corresponding DBJ accession no. and RGP clone ID.  
 A gene with identity or significant homology to a protein is  
 classified based on the protein name to indicate the homology level  
 such as same name, 'putative-' and '-like protein'. A gene without  
 significant homology to any protein but with EST homology (covering  
 almost the entire length of partial sequence) is classified as an  
 'unknown' protein. A gene predicted with a gene prediction program  
 is classified as a 'hypothetical' protein.  
 The orientation of the sequence is from -21M13 to M13rev of the BAC  
 clone. This sequence of OSJNBa0025P13 clone has an overlap with  
 P0702D12 (DBJ: AP002820) clone at the position 1 to 1534 of 5' end  
 and with P0487H02 (DBJ: AP002883) at the position 84175 to 133242  
 of 3' end. The sequence of this clone starts at the position  
 135799 of P0702D12 and ends at the position 49068 of P0487H02.  
 Detailed information on overlap and assembly quality together with  
 annotation of this entry is available at  
 http://rpg.dna.affrc.go.jp/GenomeSeq.html.  
 Location/Qualifiers  
 1..133242  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /chromosome="1"  
 /clone="OSJNBa0025P13"  
 complement(join(4221..4246,4864..5040,5244..5909,  
 6141..6189))  
 /gene="OSJNBa0025P13.1"  
 complement(join(4221..4246,4864..5040,5244..5909,  
 6141..6189))  
 /gene="OSJNBa0025P13.1"  
 /codon\_start=1  
 /product="putative protein kinase"  
 /protein\_id="BAB55742.1"  
 /db\_xref="GI:14164492"  
 /translation="MYLPMLIVLFLVLAALVLSICRIQRKKYKIHADGDOENM  
 FSVLNPFGNAVEQILEATNPFSEKICIGAGVSYVALSGKKPAVKIATENG  
 RLINQWFRYEIEATMOIRHKNIVKVGCTARDKFIYKKKGNLITAKSYSA  
 SELDMKRRICLADQVAAHLSYLDHDSDPVHEDYTKNILDERRACISDGIKTA  
 LDADSGHTRILACTGYLAPELAYTTKTEKCDVYFGVVVLELMSHEDFVSIS  
 CPKSKSTPMKDLDTRLPPAEVIALKIM"  
 join(9270..9380,9468..9560,9738..9797,9871..10143,  
 10245..10319,10435..10491,10609..10749,10829..10879,  
 10972..11067,11175..11258,11346..11448,11534..11616)  
 /gene="OSJNBa0025P13.2"  
 join(9270..9380,9468..9560,9738..9797,9871..10143,  
 10245..10319,10435..10491,10609..10749,10829..10879,  
 10972..11067,11175..11258,11346..11448,11534..11616)  
 /gene="OSJNBa0025P13.2"  
 /note="contains ESTs C72171(E1142), AU093893(E1142)"  
 /codon\_start=1  
 /product="putative shaggy-related protein kinase"  
 /protein\_id="BAB55743.1"  
 /db\_xref="GI:14164493"  
 /translation="MTSVGLIVSSSGPPEPTSGANDRLTDIISMSINDKVEAVV  
 SGSMNIGRIYTTVGRNGQPKOTISYIAERAVGSGGVFQACLTGEVAAYK  
 YLDPAKRNELDTQVLDHPNACLKHKVCGSTAAEELYNLIVNPEVETVARIRH  
 YNMSQRMPLIYKLVYQICRALAIHNCVGGCHDIRKQILVNPVHODKLCDFG  
 SAGVIVGEPNLSYICSRVYRAPELIFGATEYTAIDVAGCVLAELHLLQGVPEGD  
 SGVDQVEIIVKVGTPREIRKAMNENYEFKPOIKAPMKIKFKRMSPAVDVLS  
 RLQVSPHRCASLEVLVHPFDELADPNARLPNGRTLPLPFRFKRRELKGSAMETLV  
 KLVQVAKKQCAFLG"

## FEATURES

## source

## gene

## CDS

## CDS

## gene

gene join(15137, .15820,15928, .15963,16058, .16849)  
/gene="OSJNBa0025P13.3"  
CDS join(15137, .15820,15928, .15963,16058, .16849)  
/gene="OSJNBa0025P13.3"  
/product="putative polygalacturonase-like protein"  
/codon\_start=1  
/protein\_id="BAB55744.1"  
/db\_xref="GI:14164494"  
/translation="MELAAAGRTAATLALLALAFASSPISAADGARSARHHAKAKR  
NAHPSOAPGPAARARHAGARHGHGAPHPGRSRPARAPANPGRSFLHTT  
PAAGAATVIVDIFGAGADGVTDDTLAKTMDTACADGAGVLAAGREFLHTT  
VFPGQGSVTLQVDTIVASEPAPWANKRMVLVFRADGVSLVAGLIDGQOK  
WMDLPCPKHKGATGHCDSPVAMREPAISNNVTVMKVONBEPHFHFDNCGPH  
GISISPPALSPNTDGIHVENTSDVLTITNTVNSDDCVS IGAGTILNHIENYTCGPH  
GISISLKGAGTACVANTVRNAVITRHSNDGVRITKMOGSGSVSAVAFENRDAV  
RNPITIDQYCYLCKSCENETJAFVNGVSIGIRGVYDRGPHIRGCSDAVPCNIT  
LSDVELPASGDTVDPEFCWNTYGNATPTVPVPSGLSSGFFPYTEKKLOCY"  
/complement(join(21118, .22066,22246, .22397,23530, .23691))  
/gene="OSJNBa0025P13.4"  
/complement(join(21118, .22066,22246, .22397,23530, .23691))  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAB55745.1"  
/db\_xref="GI:14164495"  
/translation="MRAAIEQLRVQAAARAGPAGRWIRHAGADYRGGQGRSASRP  
ATGEGARATYNTKGTOTRSPYSKGVKRRLOEERKRTAAVHOGSKERLGRPAIT  
LEGIRAKTYPHPPAPPRPEVASSTTSRARSAPKPSSTGRASRANPALSRLSRO  
AGHLRRPRLSRBAGLCHRLHSVSRSSTFPNNALADATVCHALGKPEBAMP  
PTTVYVTSPPSPHQLRLCHRLIRCHRLSHGRCKSAVADAVTSHDASSR  
HCKRPAPTOPQRRRRALCRDHRGKTISSAMTRSPWPLDPLDPAVANLTATSI  
PAASINPAADAPKRRSRMPARTHSPPSPSPSPSPAPATNCRBHPHTFHR  
HHHQLLPHTDPPAAPDPDPAATARI"  
/complement(join(27109, .27305,28392, .28419))  
/gene="OSJNBa0025P13.5"  
/complement(join(27109, .27305,28392, .28419))  
/gene="OSJNBa0025P13.5"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAB55746.1"  
/db\_xref="GI:14164496"  
/translation="MDVNIIGRTNKQTANQPKKANMHSQIDTSSIGNFIDDKK  
DNLTYSYTVIAHTEYRSIYANILRLAIK"  
/complement(join(31012, .31020,31243, .31294,31358, .31478,  
32907, .33108,34245, .34337))  
/gene="OSJNBa0025P13.6"  
/complement(join(31012, .31020,31243, .31294,31358, .31478,  
32907, .33108,34245, .34337))  
/gene="OSJNBa0025P13.6"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAB55747.1"  
/db\_xref="GI:14164497"  
/translation="MHERVLISEYNDKEKIGAFHNQPLHCPGKQKEKTHPSLPVRR  
MFSKCKPAENRANQSLSPVOSHVFATVLLASVLAASGKREKSTPPDQGRGSR  
TSVATSCDCRAYIPSVVMNVILPIQLSSGMEDRLVCTPLFVDFGKQKVPNVW"  
41044, .44517  
/gene="OSJNBa0025P13.7"  
41044, .44517  
/gene="OSJNBa0025P13.7"  
/codon\_start=1  
/product="putative GAG-POL precursor"  
/protein\_id="BAB55748.1"  
/db\_xref="GI:14164498"  
/translation="MPCPVREVTEHRLAVRGAAPVQKTRQAPPERQATIREVART  
LEAGFIREVHPEWLVNVPVIVPKANGKIRKMTIDYDLNKAQPKDPYPLSRIDQVST  
AGCDLCEFLDASGVHQIRMARDEDEKTAFTPTIGTYCYTMPGLNAAPTQRTTR  
ISLSQIGRNVAAVVDLVKTRNOETLTLATPESLSARIKAMPDCVGVPGP  
KLGLVAVARGIEMANEKIRAIEMRPPSKLADVOCTGCAALSRISLGEALPL  
FKLKRSGPTWTEABERAFIOLKAYISSPVVAVAPPEDELLIYLAAPVSAALV  
VERBEDNHSANHPHASTRPGREOGGAPRPNPGAPRPTAGACGPACPTVPAPADQ  
DGPATYGRPRLSPDPEVGTGABCRGSLDERFGDAAAGDEPRKKRVGVTF  
VSEALRADKTRYPOAKMLYAILVASRKLRYFOAHRVTVVTSYPLQOILHNEGTGR

VRNALILSEPDILFEPRHAIKQALTDVPAENTPAPEPVSABEASGTSQULPHTAY  
VMPDFGSLQAGAVTLTSPSGDVLTAVLVDVFRATNMMAEYBELGLADLRVAAGG  
IRLLVIGDSQLVNVQYKEYRCSDPMDYVVRDVFRTANMEHNPGEILGRVARDNAVA  
DELSRLASRAQTPPGAFEEERLAOPSAPPLDPEVLAPEPPIGVOASGPGSAPS  
SPRLIATWTEIOTYLDLADKTLBEDREGSERPRISKRVYVGLGYLRAAGVLLKCP  
RBOGVEILADVHBEBCAHASRTIVKARQSGPYMPTLADVDVRCRACQFHR  
QTHQPAQLOTTIPUSWPFVAGDITLGFERRAEGFEYQYVADKTPKPEAPVIKI  
DKSHALFKIRIGITARFGVPPRIITNNQOTSLFEGYCKDMGIKLCEPASPAHRSNG  
OVERANAEILKSKTFTNLIKKGDSWIEKLPAVIMANRTPSRATGETSFLVYGA  
EAVPSELSTLRSPVTVYCEADODPLRDOLDLLEERRRALLRAAYOOLRRYHOR  
HVRARSLCYVDILVIRVOTRAGLSKLSPMMEGPRVIGVPRPSVRLATDGTLELPP  
WNTHLRFP"  
/complement(join(45140, .45474,45602, .46322,46382, .47023,  
47154, .48413))  
/gene="OSJNBa0025P13.8"  
/complement(join(45140, .45474,45602, .46322,46382, .47023,  
47154, .48413))  
/gene="OSJNBa0025P13.8"  
/note="hypothetical protein  
similar to Oryza sativa chromosome 3, OSJNBa0067E01.7"  
/codon\_start=1  
/protein\_id="BAB55749.1"  
/db\_xref="GI:14164499"  
/translation="MARGSAPLDGVSUPPSRIVSERQAGLPRPMPESATGRBVVAIG  
KGRPAHPYGMVSFFLSPFAMGLVPSPSTIFPMDLLEFYDLOMAHUTPMALIMTLAIPIH  
LCMEFIRVPSRLRWFMEFTVQVSPPTAVAGCTFQRGVGLNHYICVLKRDMDK  
LDEMTPLGEERARLRLSOPAPQASNRATVDLDGDVADVLRLAGLSQGLTGAMF  
GDVLRRLRPLQORBARGEYTGDEPDIYTHOGKGMWDEDFEMVQRVNLTSSVSA  
SLIQGVYPLCDPBRRAKITLIMOVASGARHGDHOGSGSRGSGSTPGGARAG  
PRDGPQGSRRADARGRKQKDGTPPPSPRGGCAVVASRRPBAVTPPOEGRRKKK  
RLRKMGIVLPKGMJSLQGVTEPLARPDVAPLPPPPRHSKGRSEAEVYTLA  
ARRBRDREARDELREAEAAQEAARVROAEARBARLRRAGSSVRAVAEARAQ

Alignment Scores:  
Pred. No.: 9 236-42 Length: 133242  
Score: 514.50 Matches: 111  
Percent Similarity: 56.50% Conservative: 28  
Best Local Similarity: 45.12% Mismatches: 60  
Query Match: 45.21% Indels: 47  
DB: Gaps: 9

US-10-021-811-36 (1-206) x AP003140 (1-133242)

QY 10 SerGlnAspProGluValAlaArgIysGlyProTyrThrMetGluGluAspLeuIleLeuMet 29  
Db 120125 GCGGAGGAGAGAGATGTTTCGAGAGGGCCGTGACGCTGAGAGAGACGCGTGTGGCC 120184

QY 30 AsnTyrTlleAlaAsnHisGlyGluGlyValTyrPheAsnSerLeuAlaIleAlaGlyLeu 49  
Db 120185 GGGTACGTGGCGCGAAGCGGAGGGCGGTGGAACGACCTGACGCGGGCGGGCGCTG 120244

QY 50 LysArgAsnGlyLysSerCysArgLeuAArgTyrPheAsnTyrLeuAspProAspValArg 69  
Db 120245 AGCGGACCGGGAGAGAGCTGCCCGCTCGGTGCTCAATACCTCGGCCACGCGGCC 120304

QY 70 ArgGlyAsnIleThrProGluGluGluLeuLeuIleMetGluLeuHisAlaIleTyrGly 89  
Db 120305 CGCGGCGACCTTCAACCCGAGGAGCAGCTGCTCATCTCGACTCCACTTCGCGTGGGGG 120364

QY 90 AsnArgTyrSerIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 109  
Db 120365 AACCGCTGTCGAGATCGCGACGACATGCCGGGAGGAGACGACAAACAGATCAAGAAC 120424

QY 110 TyrTyrArgThrArgIleGluIleLysHisIleLysGlnAlaGluAsnPheGluGluGlnIle 129  
Db 120425 TACTGCGGACCGAGGTGCGAAGACGCGCAAGCAGCTC-----GGCTCGCAC 120472

QY 130 SerAsnAsnSerGluIleAsnAsp----- 137  
Db 120473 GTCAAATAGCCGACGATCAAGACGATCATAGGACCTCTGATGCCGCGCTCTGCGAG 120532

QY 138 -----HisGlnIleSerThrSerHisValSerThrMetAlaGluProMetGluMetTyr 155

	Df	120533	AGGATCCAGCGCGCGCGCCTGTG-----TTCGAACGGCGCGCGC-----	120574
Oy	156	SerProProCys-----	TyrGlnGlyMetLeu-----	GluPro 166
		...CCCCCGTGTCGCCGCCGCGCTACTGCACAGCGGCATGCGCATCACCTGACC	...:::	...
Df	120575	---CGCCCGTGTCGCCGCCGCGCTACTGCACAGCGGCATGCGCATCACCTGACC	...	120631
Oy	167	PheSerThr-----	-GlnPhePro---	172
		:::		
Df	120632	TCAACCAACGACGTGCAGCATGCGAGGATCGTGCATGACGACGCGCAGCGCATTTCCCGTGG		120691
Oy	173	-----	ThrIleasnProAspGlnSer	179
Df	120692	TTCAGAACCAACCACTGATGACATGCGCAGGTCATCAACCGCAGCGCATTTGCTTC		120751
Oy	180	Ser-----	CysCysThrAsnAspAsnAsnIleAsnTyrTrpSer	193
		:::		
Df	120752	AAGAGCAATTGCGCGCAGCGCAGCGCACCACGACGACGAGCGTTGGATGACATTATGACATG		120811
Oy	194	MetGluAspSerTrpSer 199		
Df	120812	TTCGAGGAGGAGCTGCTG 120829		
RESULT 12				
AP002883				
LOCUS	AP002883	145491 bp	DNA	linear
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,			PLN 21-MAR-2002
ACCESSION	AP002883			
VERSION	AP002883.2 GI:15623781			
KEYWORDS				
SOURCE				
ORGANISM	Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA, clone:P0487H02. Oryza sativa (japonica cultivar-group) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Eriocarotaceae; Oryzeae; Oryza.			
REFERENCE	1 Sasaki,T., Matsumoto,T. and Yamamoto,K. Oryza sativa nipponbare (GAJ) genomic DNA, chromosome 1, PAC clone:P0487H02 Published Only in Database (2000)			
AUTHORS	2 (bases 1 to 145491) Sasaki,T., Matsumoto,T. and Yamamoto,K. Direct Submission			
TITLE	Submitted (11-OCT-2000) Takui Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program, Kamondai 2-1-2,Tsukuba, Ibaraki 305-8602, Japan (E-mail:tsasakl@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/ Tel:81-298-38-7441, Fax:81-298-38-7468)			
JOURNAL	On Sep 14, 2001 this sequence version replaced gi:10800078. Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI Nonredundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP. Protein homologies of the coding regions were searched against NCBI Nonredundant Protein database with BLASTX2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DDBJ accession no. and RGP clone ID.			
COMMENT	A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein. The orientation of the sequence is from Sp6 to T7 of the PAC clone. This sequence of P0487H02 clone has an overlap with P0682B08 (DDBJ: AB003578) clone at the position 73,961 to 145,491 of 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <a href="http://rgp.dna.affrc.go.jp/genomeseg.html">http://rgp.dna.affrc.go.jp/genomeseg.html</a> . Location/Qualifiers			
FEATURES				

gene	source
gene	1..145491 /organism="Oryza sativa (japonica cultivar-group)" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="1" /clone="P0487H02" complement(join(6. 2722..2784,4696..4740,5045..5186,5198..5407)) /gene="P0487H02.1" complement(join(6. 2722..2784,4696..4740,5045..5186,5198..5407)) /gene="P0487H02.1" /note="hypothetical protein" /codon_start=1 /protein_id="BAB67842.1" /db_xref="GI:15623782" /translation="WQSRHGKSTRTGSQGLARCSVSKDEKYLCVNCSNCPHPQTTT HCKKPSAANAPNRDARGRGRTIPRRLDRODITRPASQKGLCTGPCLRLMSQNHIT TISYVYVITVITGCLRPFVSGLSHCSGLARSGAGDGGNGVEMWHLGSGR GAGGQVNDGTGAAAAAPLPLPSPQIMPEEGGRLAAVATMGSDVPHMETSTSD DKRSKRFGLQKLAHLVPHMGSPEKYRRKSEIIPLRKEELGREDCELGRLPPI ADQRRERPREESILTPPRESILPSPADQLRRRMAADGVLYEVARIPCEIDR RMGVCSNMRVALYKLKAPAPPLPLALPESDGILPATVSCVLGCRHASPVC GARGRYGSDYDGLFLAVGGAGCQALNLKINGRQTLDTPLNAPVNSVNPARD MAIVATVSCPTGCGTACGIISSPMLVAGHYTRSIAWRMGDOVLYLVYALAE DNPLRIEVEVDLCHGAPFLFTRAEVLAACEEPYTRDSVSLPANMFLPRVH ENETVLARLVGSGKGLMVLVRLASGRQRTTSARVFOKKFNTGEEDESNRSAS FHYVSELDDELGRMLFVGLGCSYKAGDGRYAPMEGQVFLPDSTIHWIIGDAPK PYLCSNDGKMSRATLDPQGVSRCPERPGRSHSPVWILP" join(7576..7805,8020..8053) /gene="P0487H02.2" join(7576..7805,8020..8053) /gene="P0487H02.2" /note="hypothetical protein" /codon_start=1 /protein_id="BAB67843.1" /db_xref="GI:15623783" /translation="MNRPTINVLIVHEVQFLVTHVNFSSCRENNEPKISPRIQSA AVTRVAHRNKGKRAAGDGMERGVRKRWMAASNLTAHTART" join(8658..8923,8973..9006) /gene="P0487H02.3" join(8658..8923,8973..9006) /gene="P0487H02.3" /note="hypothetical protein" /codon_start=1 /protein_id="BAB67844.1" /db_xref="GI:15623784" /translation="MRAIVQGLATTHQASARTHAAVAGVVRDTTISQAVHCSNG ADWGEINKEPAPGSGKKGVLSKFTCIIRQTEVLSLSDPSSSPYIISRNV" complement(join(10181..10287,11338..11896)) /gene="P0487H02.4" complement(join(10181..10287,11338..11896)) /gene="P0487H02.4" /note="hypothetical protein" /codon_start=1 /protein_id="BAB67845.1" /db_xref="GI:15623785" /translation="MRTPTRGGAHPAGSPVCDTHAASQVWLMEKLARDVAVNG ATRISPERNTYRRLILICRGAFTTGGRQORLEAMFAGDGRPVRCRPHGARSR GTAGCSQSRITNSGDVGAGRLTGRKQMLRLSRGVCHGDDAACATRGNAAPMP CGGCPAGTGSVVLPTRPRFRMWSIVPTSNLINTKSSSSCHCKKIGKEKETIIP AAT" join(12412..12593,13512..13574,14331..14862,14954..15319) /gene="P0487H02.5" join(12412..12593,13512..13574,14331..14862,14954..15319) /gene="P0487H02.5" /note="contains EST C99297(EI0619)" /codon_start=1 /product="putative MtN21" /protein_id="BAB67846.1" /db_xref="GI:15623786" /translation="MGRSLPPLTAMVNLGFGAGNNVSKLALDTGMSPVYLAYENI IAAVLAPLVYEFRRYKSGVITKTKLVQIIFSSSIFGATINQVLVFGVSGKSTTPVA
gene	1..145491 /organism="Oryza sativa (japonica cultivar-group)" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="1" /clone="P0487H02" complement(join(6. 2722..2784,4696..4740,5045..5186,5198..5407)) /gene="P0487H02.1" complement(join(6. 2722..2784,4696..4740,5045..5186,5198..5407)) /gene="P0487H02.1" /note="hypothetical protein" /codon_start=1 /protein_id="BAB67842.1" /db_xref="GI:15623782" /translation="WQSRHGKSTRTGSQGLARCSVSKDEKYLCVNCSNCPHPQTTT HCKKPSAANAPNRDARGRGRTIPRRLDRODITRPASQKGLCTGPCLRLMSQNHIT TISYVYVITVITGCLRPFVSGLSHCSGLARSGAGDGGNGVEMWHLGSGR GAGGQVNDGTGAAAAAPLPLPSPQIMPEEGGRLAAVATMGSDVPHMETSTSD DKRSKRFGLQKLAHLVPHMGSPEKYRRKSEIIPLRKEELGREDCELGRLPPI ADQRRERPREESILTPPRESILPSPADQLRRRMAADGVLYEVARIPCEIDR RMGVCSNMRVALYKLKAPAPPLPLALPESDGILPATVSCVLGCRHASPVC GARGRYGSDYDGLFLAVGGAGCQALNLKINGRQTLDTPLNAPVNSVNPARD MAIVATVSCPTGCGTACGIISSPMLVAGHYTRSIAWRMGDOVLYLVYALAE DNPLRIEVEVDLCHGAPFLFTRAEVLAACEEPYTRDSVSLPANMFLPRVH ENETVLARLVGSGKGLMVLVRLASGRQRTTSARVFOKKFNTGEEDESNRSAS FHYVSELDDELGRMLFVGLGCSYKAGDGRYAPMEGQVFLPDSTIHWIIGDAPK PYLCSNDGKMSRATLDPQGVSRCPERPGRSHSPVWILP" join(7576..7805,8020..8053) /gene="P0487H02.2" join(7576..7805,8020..8053) /gene="P0487H02.2" /note="hypothetical protein" /codon_start=1 /protein_id="BAB67843.1" /db_xref="GI:15623783" /translation="MNRPTINVLIVHEVQFLVTHVNFSSCRENNEPKISPRIQSA AVTRVAHRNKGKRAAGDGMERGVRKRWMAASNLTAHTART" join(8658..8923,8973..9006) /gene="P0487H02.3" join(8658..8923,8973..9006) /gene="P0487H02.3" /note="hypothetical protein" /codon_start=1 /protein_id="BAB67844.1" /db_xref="GI:15623784" /translation="MRAIVQGLATTHQASARTHAAVAGVVRDTTISQAVHCSNG ADWGEINKEPAPGSGKKGVLSKFTCIIRQTEVLSLSDPSSSPYIISRNV" complement(join(10181..10287,11338..11896)) /gene="P0487H02.4" complement(join(10181..10287,11338..11896)) /gene="P0487H02.4" /note="hypothetical protein" /codon_start=1 /protein_id="BAB67845.1" /db_xref="GI:15623785" /translation="MRTPTRGGAHPAGSPVCDTHAASQVWLMEKLARDVAVNG ATRISPERNTYRRLILICRGAFTTGGRQORLEAMFAGDGRPVRCRPHGARSR GTAGCSQSRITNSGDVGAGRLTGRKQMLRLSRGVCHGDDAACATRGNAAPMP CGGCPAGTGSVVLPTRPRFRMWSIVPTSNLINTKSSSSCHCKKIGKEKETIIP AAT" join(12412..12593,13512..13574,14331..14862,14954..15319) /gene="P0487H02.5" join(12412..12593,13512..13574,14331..14862,14954..15319) /gene="P0487H02.5" /note="contains EST C99297(EI0619)" /codon_start=1 /product="putative MtN21" /protein_id="BAB67846.1" /db_xref="GI:15623786" /translation="MGRSLPPLTAMVNLGFGAGNNVSKLALDTGMSPVYLAYENI IAAVLAPLVYEFRRYKSGVITKTKLVQIIFSSSIFGATINQVLVFGVSGKSTTPVA
gene	1..145491 /organism="Oryza sativa (japonica cultivar-group)" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="1" /clone="P0487H02" complement(join(6. 2722..2784,4696..4740,5045..5186,5198..5407)) /gene="P0487H02.1" complement(join(6. 2722..2784,4696..4740,5045..5186,5198..5407)) /gene="P0487H02.1" /note="hypothetical protein" /codon_start=1 /protein_id="BAB67842.1" /db_xref="GI:15623782" /translation="WQSRHGKSTRTGSQGLARCSVSKDEKYLCVNCSNCPHPQTTT HCKKPSAANAPNRDARGRGRTIPRRLDRODITRPASQKGLCTGPCLRLMSQNHIT TISYVYVITVITGCLRPFVSGLSHCSGLARSGAGDGGNGVEMWHLGSGR GAGGQVNDGTGAAAAAPLPLPSPQIMPEEGGRLAAVATMGSDVPHMETSTSD DKRSKRFGLQKLAHLVPHMGSPEKYRRKSEIIPLRKEELGREDCELGRLPPI ADQRRERPREESILTPPRESILPSPADQLRRRMAADGVLYEVARIPCEIDR RMGVCSNMRVALYKLKAPAPPLPLALPESDGILPATVSCVLGCRHASPVC GARGRYGSDYDGLFLAVGGAGCQALNLKINGRQTLDTPLNAPVNSVNPARD MAIVATVSCPTGCGTACGIISSPMLVAGHYTRSIAWRMGDOVLYLVYALAE DNPLRIEVEVDLCHGAPFLFTRAEVLAACEEPYTRDSVSLPANMFLPRVH ENETVLARLVGSGKGLMVLVRLASGRQRTTSARVFOKKFNTGEEDESNRSAS FHYVSELDDELGRMLFVGLGCSYKAGDGRYAPMEGQVFLPDSTIHWIIGDAPK PYLCSNDGKMSRATLDPQGVSRCPERPGRSHSPVWILP" join(7576..7805,8020..8053) /gene="P0487H02.2" join(7576..7805,8020..8053) /gene="P0487H02.2" /note="hypothetical protein" /codon_start=1 /protein_id="BAB67843.1" /db_xref="GI:15623783" /translation="MNRPTINVLIVHEVQFLVTHVNFSS

gene  
CDS  
CALSTNTLPALTFMAAARFMEVSRLSAAAGQAFVGVVCGGSMIMPEYKPLRLM  
ASPIHWRPASASGAAPPAAGAAVGLVLIISCAAAWVFIOTKMSERFSAPT  
STTIIMAGVOCAGVSAAMDPAVWIKGFIDRLISVAVIYGVSGIAFALMSIO  
VGRPLFVSMPSMLLVVAIVGMAIIDEKIHGSAIGSLIYAGLYMVMWGAREMGS  
PSLDGGGGGGVVELNGKADNATLLPVFCITTHETTRNGCSN"  
/complement (join (21707, .21839,22076, .22099,22760, .22917))  
/gene="P0487H02.6"  
/complement (join (21707, .21839,22076, .22099,22760, .22917))  
/gene="P0487H02.6"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAB67847.1"  
/db\_xref="GI:15623787"  
/translation="MYREGDRLGAEAGSRAICGGEGVCAFPPIHIGDMEETGEGI  
VRADGDRKMPVRCNGMKRTHARREVSDSASACDDKRKSTKAFKFSILYLILH  
PY"  
gene  
CDS  
complement (join (23501, .23554,23930, .24028,24083, .24274))  
/gene="P0487H02.7"  
complement (join (23501, .23554,23930, .24028,24083, .24274))  
/gene="P0487H02.7"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAB67848.1"  
/db\_xref="GI:15623788"  
/translation="MMTECRRLFENTSRVGAQMRLTAPASISDGRVAXQLELELVA  
RDLOKNEIIVAGNEMQRLGLMKSFSLHQLALECGVKSIIPAHPDETSEPKLGRVQR  
RIGNVYMDEEM"  
complement (join (27604, .27686,27736, .28503,29633, .29675))  
/gene="P0487H02.8"  
complement (join (27604, .27686,27736, .28503,29633, .29675))  
/gene="P0487H02.8"  
/note="hypothetical protein  
similar to Oryza sativa chromosome 6, P0514G12.3"  
/codon\_start=1  
/protein\_id="BAB67849.1"  
/db\_xref="GI:15623789"  
/translation="MPGASWTLTAACHGTDNSANGGQCRVSWTRVCKPRDGLGLIPD  
LRVAGPALRVKMLDLSRGHPYMDLKAVERSVSDMFASTSIIGDGGSTLFWDR  
WIDRGSIASTAPDLHYVPPRFGSRVVAAGLANMSWVDCALTYVVISQPLVMD  
AVYPTQISPCVIEDLWRMTGSDCVSAOSVQAFPLGQHSFACADILMHAKKAKCF  
PLMFABORCTHADLQKRGIDSHSACPFCQLEIETNMHLLIDCVPRARQVWLGSKRG  
GHPDLYVLTSTFAASTPWYS"  
complement (join (30213, .30418,31118, .31292,32407, .32601,  
33634, .33894,33932, .34047,34149, .34246,34509, .34561))  
/gene="P0487H02.9"  
complement (join (30213, .30418,31118, .31292,32407, .32601,  
33634, .33894,33932, .34047,34149, .34246,34509, .34561))  
/gene="P0487H02.9"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAB67850.1"  
/db\_xref="GI:15623790"  
/translation="MGRHVGVLAMRSGARCPNPPRRRLIFPCRYLRSRGR  
YPSVERDSEAGRRFRFTSSSRSHRGAFAVBAAGVYASSGVLPYSSSSAPAYS  
ATTTPGTSKKATPSLGLAAGSAATCPSSSRARRRHRTTTLTAAAPAPGDESG  
GKTGMLYSHEDIRLEOHKATTELDLHEIGLVVHGRFRGAVLEWMDATTEGREA  
RREGSKNPFQBGDIGVCGGPIPCVCFRISGSDRIISDPFSLSEIYILKYH  
FKLPTMALMPBHRIPINAPKPRALLRAVELSAGRSQVTSQLAKKPLRDCVVA  
NRPAAPPHRIRIGARRLLISVAHARHVLV"  
gene  
CDS  
35906, .36787  
/gene="P0487H02.10"  
35906, .36787  
/gene="P0487H02.10"  
/codon\_start=1  
/product="putative dehydration-induced myb-related protein  
Cpm7"  
/protein\_id="BAB67851.1"  
/db\_xref="GI:15623791"  
/translation="WDAPCGWTTSAAAVSEEDVRGRPMTVEDAVLAGVVAANGEG  
RMNELARAAGLRRTGSCRLMILNLYLRADVAGDTPPEOLLILELHRMGNRSRIA  
OHMRGTUNEIKNYVRRPVOKGAKOIGDVNSRQKDMRLMLNPRIVERTHAAAS  
EBAAPPCAAAPASHGKCHSPDETSTTSSNAGSSVYTHGEFPSTNHHMTTAAVYT  
AAADWSSQCGSGSATSTSVGISYDMFGSSELLARAYDDGDASLLPDPQADTG

gene  
CDS  
DNCSWNLIEDIWSQQPY"  
complement (join (41992, .42118,43485, .43552))  
/gene="P0487H02.11"  
complement (join (41992, .42118,43485, .43552))  
/gene="P0487H02.11"  
Alignment Scores:  
Pred. No.: 1,03e-41 Length: 145491  
Score: 514.50 Matches: 111  
Percent Similarity: 56.50% Conservative: 28  
Best Local Similarity: 45.12% Mismatches: 40  
Query Match: 45.21% Indels: 67  
DB: 8 Gaps: 9  
US-10-021-811-36 (1-206) x AP002883 (1-145491)  
QY 10 SerGlnAspProGluValArglyGlyProThrMetGluGluAspLeuIleuMet 29  
Db 35951 GCGAGAGAGAGAGAGTTCGAGAGAGGCGGTGACGTGAGAGAGAGAGCGGTGCTGGCC 36010  
QY 30 AsnTyrIleAlaAsnHisGlyGlyGlyValTyrAsnSerLeuAlaAlaGlyLeu 49  
Db 36011 GGGTACGTGCGCGGACACGCGGAGGGCGGTGAAACAGCTGGCGCGGGCGGCGCTG 36070  
QY 50 LysArgAsnGlyLysSerCysAlaGluLeuArgTyrPLeuAsnTyrLeuArgProAspValArg 69  
Db 36071 AGCGCAGCGAGAGAGAGCTGCGCGCTCCGCTGCTCACTACCTCCGCCACAGCTGCGC 36130  
QY 70 ArgGlyAsnIleThrProGluGluGluLeuIleMetGlyLeuHisAlaIleAspTyrGly 89  
Db 36131 CGGCGGACTTCAACCCGAGAGAGAGCTGCTCACTGCTGACTTCACTTCCGGTGGGG 36190  
QY 90 AsnArgTyrSerLysIleAlaLysHisLeuProGlyArgTyrThrAspAsnIleLysAsn 109  
Db 36191 AACCGTGTGTCAGATCCGCGAGACATGCCGAGGAGAGACGAGACAGATCAAGAAC 36250  
QY 110 TyrTyrArgThrArgIleGlnLysHisIleLysGlnAlaGluAsnHisGlnIleGlnSer 129  
Db 36251 TACTGCGGAGCCAGGCTGAGAGAGACGCGCAAGCAGCTC-----GGCTGGAC 36298  
QY 130 SerAsnAsnSerGluIleAsnAsp----- 137  
Db 36299 GTCATATAGCCGCAAGTTCAGAGAGCTCATAGAGCACTTGATCGCGCGCTCGTCAG 36358  
QY 138 -----HisGlnLaseThrSerHisValSerThrMetAlaGluProMetGluMetTyr 155  
Db 36359 AGGATCCACCGCGCGCGCGCTG-----TGGAAACGCGCGCGCGC----- 36400  
QY 156 SerProProCys-----TyrGlnGlyMetLeu-----GluPro 166  
Db 36401 ---CCGCCGTGTCCCGCGCGCGCTGATGCACAGCGGATGTGCATCACTGACCGC 36457  
QY 167 PheSerThr-----GlnPhePro----- 172  
Db 36458 TCAACACAGACGTCGACGATGACGAGGTGTCTCGTACACGAGCGAGAGATTCCTCCGTG 36517  
QY 173 -----ThrIleAsnProAspGlnSer 179  
Db 36518 TCAGACAAACACACCTGATGACCATGGCCAGCGTCACACGACGCGACAGATGTGCTC 36577  
QY 180 Ser-----CysCysThrAsnAspAsnAsnIleAsnTyrTrpSer 193  
Db 36578 AGCAGACATGCGGACGCGAGCGGACGCGACGACGACGAGCGTGGATGATATGACATG 36637  
QY 194 MetGluAspSerTrpSer 199  
Db 36638 TTCAGAGGAGAGCTGCTG 36655  
RESULT 13 1423 bp mRNA linear PLN 30-AUG-2001  
AF262733 Arabidopsis thaliana putative transcription factor MYB108 (MYB108)  
DEFINITION mRNA, complete cds.

ACCESSION AF262733.2 GI:15375290  
 VERSION AF262733.2  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 11597504  
 2 (bases 1 to 1423)  
 Stracke, R. and Weishaar, B.  
 Direct Submission  
 Submitted (02-MAY-2000) Biochemie, Max-Planck-Institut fuer  
 Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln D-50829, Germany  
 3 (bases 1 to 1423)  
 Stracke, R. and Weishaar, B.  
 Direct Submission  
 Submitted (29-AUG-2001) Dept. Plant Breeding and Yield Physiology,  
 Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg  
 10, Koeln D-50829, Germany  
 Sequence update by submitter  
 On Aug 30, 2001 this sequence version replaced gi:8101955.  
 REMARK  
 COMMENT  
 FEATURES  
 SOURCE:  
 1..1423  
 /organism="Arabidopsis thaliana"  
 /db\_xref="taxon:3702"  
 /chromosome="3"  
 /note="ecotype: Columbia"  
 1..1423  
 /gene="MYB108"  
 1..154  
 /gene="MYB108"  
 5  
 /gene="MYB108"  
 /note="alternative transcription start site"  
 155..1126  
 /gene="MYB108"  
 /note="At3g06490; member of the R2R3-MYB family"  
 /codon\_start=1  
 /product="putative transcription factor MYB108"  
 /protein\_id="AA072668.1"  
 /db\_xref="GI:8101956"  
 /translation="MDKGRSLKNNMDEMDIKRGPMTAEPDKLANYIANGSGRM  
 NSLSRCAGLORTGKSCRLRWLNTLRDVRNGITLLEOILLIEHSRWGNRSKIAQY  
 LRGRTNENIKYWRTRVOKAKOLKCDVNSQKFDTKYLMRFLVLOSASASAA  
 AATTTTGTGAGTSCITTSNNQFMNDYNNNGGQGFVWSNNDYITPENSVAV  
 SPASDLREYISAPNPPEYSGMGNSYPDQNLVSSQLPDNYPDYSGLDLDELTLAM  
 QEOSNLSEFNISNIGASSSSDSLWNIGETDEEFLQOOOQFNNNGSF"  
 1127..1423  
 /gene="MYB108"  
 3' UTR  
 BASE COUNT 498 a 273 c 282 g 370 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.19e-43 Length: 1423  
 Score: 510.00 Matches: 106  
 Percent Similarity: 62.50% Conservative: 24  
 Best Local Similarity: 50.96% Mismatches: 42  
 Query Match: 44.82% Indels: 36  
 DB: 8 Gaps: 5  
 US-10-021-811-36 (1-206) x AF262733 (1-1423)  
 QY 1 MetAAspLysIleGlnGlnCysLysThrSerGlnAspProGlu-----ValArg 16  
 Db 155 ATGATGTAAGAAAGGAAGCTTGAAGAACACACATGGAAGACGAGATGACCTAAAG 214  
 QY 17 LrGGLyProTrpThrMetGluGluAspLeuIleLeuMetAsnTrpIleAlaAsnHISGLy 36

Db 215 AGAGTCCCGTGAAGTCTGCTGAGAAAGATTTTAAGCTCATGATTAATTAATCTCTATATGA 274  
 QY 37 GLuGlyValTrpAsnSerIleuAlaIleGlyLeuIleArgAsnGlyLysSerCys 56  
 Db 275 GAAGTCGCGTGAAGTCTCTCTCTGTCGCGCGGCTTCAACGCAACCGTAAAGCTGT 334  
 QY 57 ArgLeuArgTrpLeuAsnTrpLeuArgProAspValArgArgGlyAsnIleThrProGlu 76  
 Db 335 AGACTAAGGTGTTAAATCATCTCCGCGCTGAGCTCCGCGGAAACATTACACTTGA 394  
 QY 77 GLuGluLeuLeuIleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerIleAla 96  
 Db 395 GAACAACTCTGATCTTCACTTCACTTCCGTTGGGAAATGATGCAAAAATTCGA 454  
 QY 97 LysHisLeuProGlyArgThrAspAsnGluIleLysAsnTrpArgTrpArgIleGln 116  
 Db 455 CATATTTCACCGGAGAAAGACGACAAACGATCAAGAACTATGAGAGACCGGGTGC 514  
 QY 117 LysHisLeuIleGlnAlaIleLys-----AspPheGlnGlnSerSerAsn--- 132  
 Db 515 AACATGCCAAACAGTTGAATGTGATGTGAATGCCAACATTCAAAGACCAATGAAG 574  
 QY 133 -----SerGluLeuAspPheIleGlnAlaSerThrSerHis 144  
 Db 575 TACTTGTGATGCTCTGACTGATGAGAGATTCAGTCAAGCTCGGCTCATCCGACGA 634  
 QY 145 ValSerThrMetAlaGluProMetGluMetLysSerProProCysTrpGlnGlyMetLeu 164  
 Db 635 GCAGCCACACAC----- 646  
 QY 165 GluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSerCysThrAsn 184  
 Db 647 -----ACAACACACACACACAGATCAGCGGACGCTATCTTGATCACAAC 697  
 QY 185 AspAsnAsnAsn---IleAsnTrp 191  
 Db 698 TCTAACATCATTCATCATGATTAC 721  
 RESULT 14  
 AF334815 1140 bp mRNA linear PLN 30-AUG-2001  
 LOCUS Arabidopsis thaliana putative transcription factor (MYB116) mRNA,  
 DEFINITION complete cds.  
 ACCESSION AF334815  
 VERSION AF334815.2 GI:15375294  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 11597504  
 2 (bases 1 to 1140)  
 Stracke, R. and Weishaar, B.  
 Direct Submission  
 Submitted (10-JUN-2001) Dept. Plant Breeding and Yield Physiology,  
 Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg  
 10, Koeln D-50829, Germany  
 3 (bases 1 to 1140)  
 Stracke, R. and Weishaar, B.  
 Direct Submission  
 Submitted (29-AUG-2001) Dept. Plant Breeding and Yield Physiology,  
 Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg  
 10, Koeln D-50829, Germany  
 Sequence update by submitter  
 On Aug 30, 2001 this sequence version replaced gi:13430155.  
 REMARK  
 COMMENT  
 FEATURES  
 LOCATION/Qualifiers



```

source      1. .1140
             /organism="Arabidopsis thaliana"
             /db_xref="taxon:3702"
             /note="ecotype: Columbia"
gene         1. .1140
             /gene="MYB116"
5' UTR       1. .129
             /gene="MYB116"
             /note="MYB116"
             /note="At1g25340; R2R3-MYB family member"
             /codon_start=1
             /product="putative transcription factor MYB116"
             /protein_id="AAK25748.2"
             /db_xref="GI:15375295"
             /translation="MSNITKKKCNKNEBGAQKGPWTLSEPTLLNNYISHGEGRN
             LKASGLKRAKSCRLRWLNTYLPDIKRNLTLPQDLLELHSMKRNKSTKYL
             PGRNDNDIKYWRTRVQKQKQNLIDNSHRTLEVRSPFPELLEINIDKNSNNIK
             ANAPDLGPIRLDSKIDGFNNMDCSTMSMSDLKTSQFMDPSLETITMSLEGRGSS
             OCSEVYSSFPCLSEEEYMAVMGSSDISALHDCHVADSKEYEDVTDLMNMMDIWQF
             NEYAHFN"
             982. .1140
             /gene="MYB116"
BASE COUNT  397 a 202 c 222 g 319 t
ORIGIN
Alignment Scores:
Pred. No.: 2,06e-42 Length: 1140
Score: 497.00 Matches: 96
Percent Similarity: 65.52% Conservative: 37
Best Local Similarity: 47.29% Mismatches: 44
Query Match: 43.67% Indels: 26
              Gaps: 4
US-10-021-811-36 (1-206) x AF334815 (1-1140)
Qy 4 LysGInGInCyLyThrSergInAsp---ProGluValArgLyGlyProTirPThrMet 22
Db 145 AAGAAGAGTGAATGAAATGAAGAGGTCAGAGCAGAGAAAGGCTTGAGACACTC 204
Qy 23 GUGUUAAPLLEULEUMETASNTYRILEALAAENHISGLYGLUGLYVALTRPAsnSer 42
Db 205 GAGGAAGACACTCTTCACCAATTACATTTCCATTAACGGTGAAGCCGATGGAATCTG 264
Qy 43 LeuAlaLysAlaIaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTirPLeuAsn 62
Db 265 CTCGCTTAATCTTCTGGGCTTAAGAGACAGAGAAAAGTTGATGATTAAGTGTGAAT 324
Qy 63 TyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeuMet 82
Db 325 TACCTTAATCCGACATTAAGCGTGGGAATCTCACTCCTCAAGAACAACTTTTAATCCTT 384
Qy 83 GULeuHISaLaLyTTCGLyAsnArgTirPserLySILEaLyHisLeuProGlyValArg 102
Db 385 GAGCTCCATTTCAATAGGGGTAATAGGTGTCAAAATTTCCAAATATTTACCAAGAAAGA 444
Qy 103 ThrAspAsnGluLileLysAsnTyrTirPArgThrArgIleGlnLysHisIleLysGlnAla 122
Db 445 ACGACCAACGAAATATCAAAACCTACTGAGAACTAGAGTCCAGAAACAAACGACGACGCTC 504
Qy 123 GluAsnPhgInGlnInSerSerAsnAsnSergLulleAsnAspHisGlnAlaSerThr 142
Db 505 -----AAACATAGATTCACATAGCCACACAAGTTC----- 531
Qy 143 SerHisValSerThrMetAlaGluProMetGluMetYrSerProCysArgTyrGlnGly 162
Db 532 -----ATAGAAAGTTGTTCCGTAGCTTTTGGTTTCCAAAGA 564
Qy 163 MetLeuGluProPhaSerThrGlnPheProThrIleAsn-----ProAsp 177
Db 565 CTGATCAACAGAGATTAAACAACTACATACCAACAAATATTAAGCTAATGCTCTGAT 624
Qy 178 GlnSerSerCysCysThrAsnAspAsnAsnIleAsnTyrTirPserMetGluAspSer 197

```

```

Db 625 TTACTTGACCAATTTCAGAGACAGCAAGATTGGGTTTCAACACATGATGATTGTC 684
Qy 198 TirPserMet 200
Db 685 ACTTCCATG 693
RESULT 15
AY008377 931 bp mRNA linear PLN 30-AUG-2001
LOCUS AY008377
DEFINITION Arabidopsis thaliana putative transcription factor MYB112 mRNA,
complete cds.
ACCESSION AY008377
VERSION AY008377.2 GI:15375307
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
AUTHORS Stracke,R., Werber,M. and Weisshaar,B.
TITLE The R2R3-MYB gene family in Arabidopsis thaliana
JOURNAL Curr. Opin. Plant Biol. 4 (5), 447-456 (2001)
MEDLINE 21481677
PUBMED 11597504
REFERENCE
AUTHORS Stracke,R. and Weisshaar,B.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-2000) Dept. Biochemie, Max-Planck-Institut fuer
Zuechtungsforshung, Carl-von-Linne-Weg 10, Koeln D-50829, Germany
3 (bases 1 to 931)
REFERENCE
AUTHORS Stracke,R. and Weisshaar,B.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2001) Dept. Plant Breeding and Yield Physiology,
Max-Planck-Institut fuer Zuechtungsforshung, Carl-von-Linne-Weg
10, Koeln D-50829, Germany
Sequence update by submitter
On Aug 30, 2001 this sequence version replaced gi:11641119.
FEATURES
source
1. .931
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/note="ecotype: Columbia"
1. .83
/note="alternative transcription start site"
6
/note="84. .815
/codon_start=1
/product="putative transcription factor MYB112"
/protein_id="AAK38379.2"
/db_xref="GI:15375308"
/translation="MNSITRKPNKTKTLINKHEVBEVKEKMEIRIRGPWTVEDMK
LVSYSILHGEGRWNSLSRSAGLNRGKSCRLRWLNTYLPDIRGDISLOEFTIIEHL
SPWGRWSKIACHLPGRDNEIKYWRTRVQKQKQNLIDNSHRTLEVRSPFPELLEINIDKNSNNIK
IRIATOSVQFTSNHSYSPENSVAITATSSSSSEAVSSFPYGQVQVEFGTLLDHNTNG
GYFNGGDFFETLCSFDELKMKVLIO"
816. .931
3' UTR
897
polya_site
/note="alternative"
BASE COUNT 317 a 183 c 194 g 237 t
ORIGIN
Alignment Scores:
Pred. No.: 2,89e-41 Length: 931
Score: 485.00 Matches: 90
Percent Similarity: 69.02% Conservative: 37
Best Local Similarity: 48.91% Mismatches: 39
Query Match: 42.62% Indels: 18
              Gaps: 4

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 8, 2003, 10:30:00 ; Search time 223 Seconds

(Without alignments)  
2080.321 Million cell updates/sec

Title:

US-10-021-811-36

Perfect score:

1138

1 MDKQOQCKTSQDEVRKGPW.....NNINYSMEDSMQMLNGD 206

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched:

2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters:

4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODL=frame+ p2n.model -DEV=xlh  
-Q=/cgm\_1/USPFO.spool/US10021811/rnatat\_03022003\_11221\_25629/app\_query.fasta.1.391  
-DB=N Geneseq\_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALITG=200 -THR SCORE=DCT -THR MAX=100 -THR MTN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pcio -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USR=US10021811.cgm.1.1.187 @rnatat\_03022003\_11221\_25629 -NCPU=6 -ICPU=3  
-NO XLPXY -NO MMAR -LARGEOUTRY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WAP TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq\_101002: \*  
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: \*  
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: \*  
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: \*  
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: \*  
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: \*  
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: \*  
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: \*  
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: \*  
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: \*  
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: \*  
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: \*  
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: \*  
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: \*  
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: \*  
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: \*  
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: \*  
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: \*  
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: \*  
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: \*  
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: \*  
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: \*  
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT: \*  
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: \*  
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length	ID
1	759.5	66.7	862	21 AAC57336
2	671	59.0	910	21 AAC40587
3	667	58.6	908	21 AAC48506
4	625.5	55.0	907	21 AAC57337
5	585	51.4	621	21 AAC41714
6	585	51.4	1127	21 AAC37953
7	540	47.5	673	21 AAC36197
8	539	47.4	1137	22 AAD05784
9	523.5	46.0	981	24 ABR65252
10	510	44.8	972	24 ABR65364
11	495	43.5	626	22 AAH87724
12	478	42.0	524	21 AAC57194
13	475	41.7	837	20 AAX25572
14	459	40.3	959	21 AAC44894
15	421	37.0	389	21 AAC56152
16	421	37.0	417	21 AAC56774
17	410.5	36.1	516	21 AAC56352
18	410.5	36.1	516	21 AAC56457
19	406.5	35.7	1384	21 AAC42228
20	406	35.7	660	24 ABR65183
21	405.5	35.6	542	21 AAC56434
22	404.5	35.5	389	21 AAC56837
23	404	35.5	1081	22 AAF90596
24	402.5	35.4	1820	21 AAZ46048
25	402.5	35.4	1820	22 AAH28336
26	400.5	35.2	774	21 AAC42662
27	397.5	34.9	1209	22 AAD05748
28	397.5	34.9	1209	24 ABR65181
29	396.5	34.8	1006	22 AAF90594
30	396	34.8	2220	18 AAT60744
31	392	34.4	1218	21 AAC57346
32	391.5	34.4	612	24 ABR65317
33	390.5	34.3	800	24 ABR65247
34	390	34.3	526	21 AAC57245
35	387	34.0	1369	22 AAD05764
36	386.5	34.0	1825	21 AAC44207
37	386	33.9	2352	18 AAT60745
38	385	33.8	1062	21 AAC43611
39	385	33.8	1087	22 AAD06456
40	384	33.7	1081	24 ABR65248
41	383.5	33.7	949	22 AAF85188
42	383.5	33.7	2521	20 AAX25571
43	383	33.7	573	21 AAC57325
44	382.5	33.6	843	21 AAC43054
45	382.5	33.6	1044	24 ABR98399
ALIGNMENTS				
RESULT 1				
AAC57336	ID	AAC57336 standard; DNA; 862 BP.		
XX	25-JAN-2001	(first entry)		
XX	Eucalyptus grandis transcription factor DNA sequence #773.			
XX	Plant; transcription factor; gene expression; eucalyptus; pine; acacia;			
XX	poplar; sweetgum; teak; mahogany; bzlp; G-box binding factor;			
XX	basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;			
XX	homeodomain zipper; LIM domain; AP2; ERBS; zinc finger domain;			
XX	type 2 Cys2His2; CCAAT box element; MYB; ss.			
XX	Eucalyptus grandis.			

PN WO200053724-A2.  
XX  
XX 14-SEP-2000.  
XX  
XX 09-MAR-2000; 2000WO-US06112.  
XX  
XX 11-MAR-1999; 99US-0266513.  
XX 18-AUG-1999; 99US-0149485.  
XX  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
XX  
XX Wood M, McGrath A, Shenk MA, Glenn M;  
XX WPI; 2000-579369/54.  
XX  
XX New isolated polynucleotide encoding a plant transcription factor for  
XX producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
XX having modified gene expression or modified activity of a polypeptide  
XX -  
XX  
XX Claim 1; Page 649; 747pp; English.  
XX  
XX The present invention relates to novel plant transcription factors from  
XX Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
XX sequence for one such transcription factor. The transcription factor may  
XX be used to produce a plant having modified gene expression such as a  
XX woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
XX mahogany species or to modify the activity of a polypeptide in a plant.  
XX The transcription factors of the present invention are members from the  
XX following families of regulatory proteins: bZIP, bZIP family of G-box  
XX binding factors, basic helix-loop-helix zipper, LIM domain, AP2  
XX homeoic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
XX and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
XX  
XX SQ Sequence 862 BP; 268 A; 187 C; 213 G; 194 T; 0 other;  
XX  
XX Alignment Scores:  
XX Pred. No.: 1,7e-74 Length: 862  
XX Score: 759.50 Matches: 152  
XX Percent Similarity: 76.30% Conservative: 9  
XX Best Local Similarity: 72.04% Mismatches: 29  
XX Query Match: 66.74% Indels: 21  
XX DB: Gaps: 5  
XX  
XX US-10-021-811-36 (1-206) x AAC57336 (1-862)  
XX  
XX QY 1 MetAspIysLys-----GInGInCySerThrSerGInAspProGluValArgLysGly 18  
XX Db 31 ATGGACAAAGAGCCAGACGACGAGTGAAGTCCCAAGATGTCGAGGAGGAAAAAGCG 90  
XX QY 19 ProTPrThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGly 38  
XX Db 91 CCGTGGACGATGGAAGAGATCTCATCTCATCAATACATAGGATACCGGCGAAGGC 150  
XX QY 39 ValTrpAsnSerIleuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeu 58  
XX Db 151 AATTGGAACTCCCTTAGCCAAAGCTGCTGTAAAGCTACCGGAAAGAGATTCTCGCTC 210  
XX QY 59 ArgTPrLeuAsnTyrLeuAspProAspValArgArgGlyAsnIleThrProGluGluGln 78  
XX Db 211 CAGTGGCTGAATCTATCGCAACCGACGTCGGAGAGCAACATCACTAGTGAGGACG 270  
XX QY 79 LeuLeuIleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerIysIleAlaLysHis 98  
XX Db 271 CTCTGTATCATGAACTGCATCCCAAGTGGGAAACAGGGTGGCTTAAATTTGCAAGCAT 330  
XX QY 99 LeuProGlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHis 118  
XX Db 331 CTTCCTCCGGAAGACTGCAATGAGATTAAGAACTTCTGAGGAGACTGAATCCAAAAGCAC 390  
XX QY 119 I.LelysGlnIlaGluAsnPheGlnGlnInSerSerAsnAsnSerGluIleAsnAspHis 138

Db 391 ATCAGCAAGCAGAGAGGCTTCTCTGTCAGAGCTCCGAGATAGTCAT----- 438  
XX  
XX QY 139 GlnAlaSerThrSerHisValSerThrMetAlaGluProMetGluMetTyr---SerPro 157  
XX Db 439 CAAGCAGAGCAAGACCAATGTCAGCATGCCAGCCATGAGACCTACGATCACCAG 498  
XX QY 158 ProCysTyrGlnGly-----MetLeuGluProPheSerThrGlnPheProThrIleAsn 175  
XX Db 499 CCGTCAATCCAAAGGCAACAACAATGAGAGCTTTGCCGTCGAATTGTGGTC----- 552  
XX QY 176 ProAspGlnSerSerCysCysThrAsnAspAsnAsnIleAsnTyrTrpSerMetGlu 195  
XX Db 553 -----GAGTCAAATGAAGCCTACTGAGCATGAGAC 582  
XX  
XX QY 196 AspSerTrpSerMetGluLeuLeuAsnGlyAsp 206  
XX Db 583 GATCTTGGTCTATGACAGTTACTCATGCGGAT 615  
XX  
XX RESULT 2  
XX AAC40587  
XX ID AAC40587 standard; DNA; 910 BP.  
XX  
XX AC AAC40587;  
XX  
XX 17-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 28847.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway;  
XX metabolic pathway; promoter; termination sequence; ss.  
XX  
XX Arabidopsis thaliana.  
XX  
XX OS  
XX  
XX PN EP1033405-A2.  
XX  
XX PD 06-SEP-2000.  
XX  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX XX 25-FEB-1999; 99US-0121825.  
XX XX 05-MAR-1999; 99US-0123180.  
XX XX 09-MAR-1999; 99US-0123548.  
XX XX 23-MAR-1999; 99US-0125788.  
XX XX 25-MAR-1999; 99US-0126264.  
XX XX 29-MAR-1999; 99US-0126785.  
XX XX 01-APR-1999; 99US-0127462.  
XX XX 06-APR-1999; 99US-0128234.  
XX XX 08-APR-1999; 99US-0128714.  
XX XX 16-APR-1999; 99US-0129845.  
XX XX 19-APR-1999; 99US-0130077.  
XX XX 21-APR-1999; 99US-0130449.  
XX XX 23-APR-1999; 99US-0130510.  
XX XX 23-APR-1999; 99US-0130891.  
XX XX 28-APR-1999; 99US-0131449.  
XX XX 30-APR-1999; 99US-0132048.  
XX XX 30-APR-1999; 99US-0132407.  
XX XX 04-MAY-1999; 99US-0132484.  
XX XX 05-MAY-1999; 99US-0132485.  
XX XX 06-MAY-1999; 99US-0132486.  
XX XX 06-MAY-1999; 99US-0132487.  
XX XX 07-MAY-1999; 99US-0132863.  
XX XX 11-MAY-1999; 99US-0134256.  
XX XX 14-MAY-1999; 99US-0134258.  
XX XX 14-MAY-1999; 99US-0134219.  
XX XX 14-MAY-1999; 99US-0134221.  
XX XX 14-MAY-1999; 99US-0134370.  
XX XX 18-MAY-1999; 99US-0134768.  
XX XX 19-MAY-1999; 99US-0134941.  
XX XX 20-MAY-1999; 99US-0135124.  
XX XX 21-MAY-1999; 99US-0135353.  
XX XX 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139889.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142350.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144088.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 22-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0145386.  
PR 02-AUG-1999; 99US-0145388.  
PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149425.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.

PR 28-OC<sup>1</sup>-1999; 99US-0161993.  
PR 29-OC<sup>1</sup>-1999; 99US-0162142.

## Alignment Scores:

Pred. No.:	1,16e-64	Length:	910
Score:	671.00	Matches:	133
Percent Similarity:	72.90%	Conservative:	23
Best Local Similarity:	62.15%	Mismatches:	50
Query Match:	58.96%	Indels:	8
DB:	21	Gaps:	4

US-10-021-811-36 (1-206) x AAC40587 (1-910)

```
QY 1 MetAspLysGlnGlnGlnCys-----LysThrSerGlnAspProGluValArgLysGly 18
    |||:::|||||
DB 110 ATGGAGAAAGAGAAAGTAGTGTGTTCTGGATCAGAGATGACAGAGTGAAGG 169
QY 19 ProTPThrMetGluGluAspLeuIleLeuMetAsnYrTleAlaAsnHisGlyGluGly 38
    |||:::|||||
DB 170 CCATGGACGATGAGAAAGATTGATTTCTCATCAATTATATGCCAATCATGTGGAAGGT 229
QY 39 ValTrpAsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeu 58
    |||:::|||||
DB 230 GTTGGAACTCTCTCGCAAACTCGCGAGCTAAACGCCGGGAAAGTTGCCGCTC 289
QY 59 ArgTrpLeuAsnYrTleuArgProAspValArgArgGlyAsnIleThrProGluGln 78
    |||:::|||||
DB 290 CGGTGGCTGAACCTACCTCCGACCTGATGTGTGCGACGGGAAATATCACACAGAAACAG 349
QY 79 LeuLeuIleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHis 98
    |||:::|||||
DB 350 CTCACCATCAATGAACTTCATGCAAAATGGGAAATAGTGTCTAAATTCGCAAAACAT 409
QY 99 LeuProGluArgTrpAspAsnGluIleLysAsnYrTTrpArgTrpArgIleGlnLysHis 118
    |||:::|||||
DB 410 TTACCAAGAAAGACCGCAATGATGAATAAGAACTTTGGAGATAGATCCGAAATAC 469
QY 119 IleLysGlnIaGluAsnPhenGlnGlnSerSerAsnSerGluIleAsnAspHis 138
    |||:::|||||
DB 470 ATTCATCAAGACGAGAAACGACGACCGTTGATCACAAAAGCTCCGATTATAAACAT 529
QY 139 GlnAlaSerThrSerHisVal---SerThrMetAlaGluProMetGluMetYrSerPro 157
    |||:::|||||
DB 530 CATGCGCAACGAGCCATGTCATGATGATGATCAAGAAACCATGATGATTTCCCA 589
QY 158 ProCysTrpGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsn----- 175
    |||:::|||||
DB 590 ACGACGTCGATCAACATGCCAGCAATATTATATGAGCTTAATTATGATG 649
QY 176 ProAspGlnSerSerCys-----CysThrAsnAspAsnAsnIleAsnYrTrp 192
    |||:::|||||
DB 650 CTTAAATCCAGTTGATCATGATGCCATTATCTGTTGATCATCCGAAACAAACTATTGG 709
QY 193 SerMetGluAspSerTrpSerMetGlnLeuLeuAsnGlyAsp 206
    |||:::|||||
DB 710 AGCGTCGATGATCTTTGGCCCATGAATATATATATATGTAAT 751
RESULT 3
ID AAC48506 standard; DNA; 908 BP.
AC AAC48506;
XX AAC48506;
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 57727.
XX
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
OS
XX
XX
```

```
PN EP1033405-A2.
PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129846.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134758.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137232.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139482.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 22-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 26-JUN-1999; 99US-0140921.
PR 29-JUN-1999; 99US-0141287.
PR 30-JUN-1999; 99US-0141842.
PR 01-JUL-1999;
```

PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143324.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-01455913.  
PR 27-JUL-1999; 99US-0145919.  
PR 27-JUL-1999; 99US-0145951.  
PR 28-JUL-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159295.  
PR 13-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 22-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:  
Pred. No.: 3,22e-64 Length: 908  
Score: 667.00 Matches: 132  
Percent Similarity: 72.43% Conservative: 23  
Best Local Similarity: 61.68% Mismatches: 51  
Query Match: 58.61% Indels: 8  
DB: 21 Gaps: 4

US-10-021-811-36 (1-206) x AAC48506 (1-908)

QY 1 MetAspLysIleGlnGlnCys-----LysThrSerGlnAspProGluValArgLysGly 18  
DB 108 ATGAGAAAGAAAGAAAGTGTGTGCTGTGATCAGAGATCGAGAGGTGAGAAAGG 167  
QY 19 ProTyrThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGly 38  
DB 168 CCATGAGCATGGAAGAAAGATTGATTCATCAATTATATGCGCAATCATGTGAAGGT 227  
QY 39 ValTyrAsnSerIleuAlaLysAlaIleGlyLeuLysArgAsnGlyLysSerCysArgLeu 58  
DB 228 GTTTGAAACTCTCTCCCAAACTTCGAGACTTAAACGACCGGAGAAAGTTGCCGCTC 287  
QY 59 ArgTyrLeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGln 78  
DB 288 CGGTGGCTGACTACTCTCCGACTGATGTGCGACGGGAAATATCACACAGAGAGACG 347  
QY 79 LeuLeuIleMetGluLeuHisAlaLysTyrGlyAsnArgTyrSerLysIleAlaLysHis 98  
DB 348 CTCACCATCATGGAAGCTTCATGCAAAATGGGGAAATAGGTGTGTCAAAATTCCAAGCAT 407  
QY 99 LeuProGlyValArgThrAspAsnGluIleLysAsnTyrTyrArgIleGlnLysHis 118





KM Hybridisation assay; genetic mapping; gene expression control;  
KM protein identification; signal transduction pathway;  
KM metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX EPI033405-A2.  
PN  
XX 06-SEP-2000.  
PD  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132464.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145951.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.

```
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155466.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.
```

## Alignment Scores:

```
Pred. No.: 2.29e-55 Length: 621
Score: 585.00 Matches: 121
Percent Similarity: 65.14% Conservative: 21
Best Local Similarity: 55.50% Mismatches: 44
Query Match: 51.41% Indels: 32
DB: 21 Gaps: 6
```

US-10-021-811-36 (1-206) x AAC41714 (1-621)

```
QY 1 MetAspLysLeuGlnGlnCysLys-----ThisSerGlnAspProGlu----- 14
Db 13 ATGGAAGAGAAAGGAGGAGGAGCAATTAACGACAGAAAGAAAGAAAGAAACA 72
QY 15 ValArgLysGlyProTPThrMetGlnGluAspLeuLeuMetAsnTyrlleAlaasn 34
Db 73 GTGAGAAAGAGACCTTGACATGAGAAAGATTTCATCTCTTTAAATTACATCTTAAT 132
QY 35 HisGlyGlnGlyValITPAsnSerLeuAlaLysAlaAlaGlyLeuLysSarGAsnGlyLys 54
Db 133 CAGGTATAGAGTCTTTGAACTGTGCGCAAAAGCCTGTGTTAAACCTACCTGAAAAA 192
QY 55 SarCysArgLeuArgTrrPLeuAsnTyrlleuArgProAspValArgArgGlyAsnIleThr 74
Db 193 AATTGTGGCTCGGTGGTGAATCTATCTCCGACCAAGATGTGGGGAGGAAACATTAAC 252
```

```
QY 75 ProGlnGlnLeuLeuIleMetGluLeuHisAlaLysTrrGlyAsnArgTrrSerLys 94
Db 253 GCAGAGAAACACGCTTTGTATCATTCAGCTTCATGCTTAAGCTTGAAACAGGGTGCAGAG 312
QY 95 IleAlaLysHisLeuProGlyArgThrAspAsnGlnIleLysAsnTyrlleArgThrArg 114
Db 313 ATTGGAGACATCTTCCGGGAAAGAAAGGACGACGATTAAGAACTTTCGAGAGCAAGAG 372
QY 115 IleGlnLysHisIleLysGlnAlaGlu-----AspHeGlnGlnIleSer 130
Db 373 ATTCAGAGACACATGATAAGTGCATCGGAAATATATGATGATCATCATCATTTGTTCC 432
QY 131 AsnAsnSerGlnIleAsn-----AspHisGlnAlaSerThrSerHisValSerThrMet 148
Db 433 GGAAACTCAGACAGCTCCGGGATGACGACGACGACGACGACGACGACGACGACGACGAC 492
QY 149 AlaGluProMetGluMetLysSerProProCysTyrlleGlnGlyMetLeuGluProPheSer 168
Db 493 GCTGAGAGCTTCTCT-----CAGGCGAAGACG 519
QY 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysThrAsnAspAsnAsn 188
Db 520 ACGACGTTTAATGTGGTG-----GAACAACAGTCAAAC 552
QY 189 IleAsnTyrlleTrrSerMetGluAspSerTrrSerMetGlnLeuAsnGlyAsp 206
Db 553 GAGATTACTGAGACGTTGAGATCTGTGGCCCGTCCACTTCTTAATGTGTAC 606
RESULT 6
AAC37953
ID AAC37953 standard; DNA; 1127 BP.
XX
AC AAC37953;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 19258.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
XX
PR 23-MAR-1999; 99US-0125788.
XX
PR 25-MAR-1999; 99US-0126284.
XX
PR 29-MAR-1999; 99US-0126785.
XX
PR 01-APR-1999; 99US-0127462.
XX
PR 06-APR-1999; 99US-0128714.
XX
PR 08-APR-1999; 99US-0129845.
XX
PR 16-APR-1999; 99US-0130077.
XX
PR 19-APR-1999; 99US-0130077.
XX
PR 21-APR-1999; 99US-0130449.
XX
PR 23-APR-1999; 99US-0130510.
XX
PR 23-APR-1999; 99US-0130891.
XX
PR 28-APR-1999; 99US-0131449.
XX
PR 30-APR-1999; 99US-0132048.
XX
PR 30-APR-1999; 99US-0132407.
XX
PR 04-MAY-1999; 99US-0132484.
XX
PR 05-MAY-1999; 99US-0132485.
XX
PR 06-MAY-1999; 99US-0132486.
XX
PR 07-MAY-1999; 99US-0132487.
XX
PR 11-MAY-1999; 99US-0132863.
XX
PR 11-MAY-1999; 99US-0134256.
```





Db 182 GAGAGCTCTTGATCTCGAGCTCCATTCGCCGCGCAATGATGTCGAAGATCGG 241  
Qy 97 LysHisIleuProGlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGln 116  
Db 242 CAACACTTGGCGGGGAGGAGCGACAGACAGATCAAGAACTACTGGGCAACCGGGTGGG 301  
Qy 117 LysHisIleLysGlnIleGluAsnPhenGlnGlnSerSerAsnAsnSerGluIleAsn 136  
Db 302 AAGCAGCGGAGGAGCTC-----AAATGTGACTCAACAGCAGATTCAAG 349  
Qy 137 AsnHis-----GlnAsnSerThrSer 143  
Db 350 GAGCGCATGAATACTCTGTGATGCCAGGCTGTGAGAGATCCAGCGCTCGCC 409  
Qy 144 HisValSerThr-----MetIleGluProMetGluMetTyrSer 156  
Db 410 TCTGTCTCGACCGTACTGTGCGCGCGCGCCATGGCACCACCAACAGATGGCCACC 469  
Qy 157 -----ProProCysTyrGlnGlyMetLeuGlu 165  
Db 470 ACCGACAGATCCAAATCGCGGCGATGCTTCCCGCGCCCTGGCGGCGATGGCGGC 529  
Qy 166 ProPheSer-----ThrGlnPheProThrIleAsnProAspGlnSerSerCys 181  
Db 530 GACTTCAGGGGCGGCGGAGTGAATGTGGCGCCAGCTACAGACCCCGGAGAACTCTGC 589  
Qy 182 CysThrAsnAspAsnAsnAsnIleAsnTyrTrpSerMetGluAspSerTyrSerMetGln 201  
Db 590 ACGACG-----GCGTCTTCGACCTAATTCGGTGGCGGAC 622  
Qy 202 Leu 202  
Db 623 GTC 625

## RESULT 8

AAD05784 ID AAD05784 standard; cDNA; 1137 BP.

AC AAD05784;

DT 31-JUL-2001 (first entry)

DE Arabidopsis thaliana transcription factor, G1324 cDNA.

XX Transcription factor; biochemical characteristic; controlling element;  
XX structural characteristic; developmental characteristic; gene therapy;  
XX agricultural biotechnology; plant trait modification; ss.

OS Arabidopsis thaliana.

FH Key Location/Qualifiers

FT CDS 54..914 /tag= a /product= "Transcription factor, G1324"

PN W0200136597-A1.

PD 25-MAY-2001.

PF 14-NOV-2000; 2000WO-US31344.

PR 17-NOV-1999; 99US-0166228.

PR 17-APR-2000; 2000US-0197899.

PR 22-AUG-2000; 2000US-0227439.

PA (MEND-) MENDEL BIOTECHNOLOGY INC.

PA (YUGG/) CREELMAN R.

PA (ADAM/) ADAM L.

PA (RIEC/) RIECHMANN J L.

PA (HEAR/) HEARD J.

PA (SAMA/) SAMAH R.

PA (PLUG/) PILGRIM M.

PA (PINE/) PINEDA O.  
PA (JIANG/) JIANG C.  
XX Creelman R, Yu G, Adam L, Riechmann JL, Heard J, Samaha R,  
PI Pilgrim M, Pineda O, Jiang C;  
XX WPI, 2001-335999/35.  
DR P-PSDB; AAE01900.

PT Nucleic acids encoding plant transcription factor polypeptides, useful  
PT for altering the biochemical characteristics of plants e.g. corn,  
PT potato and cotton plants -

PS Claim 4; Page 87-88; 127pp; English.

XX The present sequence is Arabidopsis thaliana transcription factor,  
CC G1324 cDNA. The transcription factor is used for altering a plant's  
CC biochemical characteristics. The transcription factor may be used to  
CC alter the structure and developmental characteristics of plants such as  
CC soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower,  
CC alfalfa, sugar cane, turf, banana, blackberry, blueberry, strawberry,  
CC raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant,  
CC grapes, honey dew, lettuce, mango, melon, onion, papaya, peas, peppers,  
CC pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon,  
CC roseaceous fruits and/or vegetable brassicas. Transcription factors are  
CC key controlling elements of biological pathways and altering expression  
CC levels of 1 or more transcription factors can change entire biological  
CC pathways in an organism. Therefore manipulating transcription factor  
CC levels in plants offers great potential in agricultural biotechnology  
CC for modifying a plant's traits. Transcription factor cDNA is useful in  
CC gene therapy.

SQ Sequence 1137 BP; 378 A; 213 C; 233 G; 313 T; 0 other;

## Alignment Scores:

Pred. No.:	6,63e-50	Length:	1137
Score:	539.00	Matches:	107
Percent Similarity:	66.18%	Conservative:	28
Best Local Similarity:	52.45%	Mismatches:	43
Query Match:	47.36%	Indels:	26
DB:	22	Gaps:	5

US-10-021-811-36 (1-206) x AAD05784 (1-1137)

Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTyr 20  
Db 66 ATGAAAGAAAGAAAGAGCTTCAAGAAAGTGAAGAACTAAGAAAGAGGCTTGG 125  
Qy 21 ThrMetGluGluAspLeuIleuMetCysTyrIleAlaAsnHisGlyGluGlyValTyr 40  
Db 126 ACTTGGAGGAAAGACACACTTCTCAAAATTACATCTCCATTAACGAGGTGGTGG 185  
Qy 41 AsnSerLeuAlaLysAlaIleGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTyr 60  
Db 186 AATCAGCTGCGCAAAATGCTGCTGCTAAAGAACTGGGAAAAGTTGATGAGATGG 245  
Qy 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGlnGlnLeu 80  
Db 246 TTGAATTACTTGAACCAGCAATAGAGAGGAGATCTTACTCTCAGAAACAGCTTTG 305  
Qy 81 IleMetGluLeuHisAlaLysTyrGlyAsnArgTyrSerLysIleAlaLysHisLeuPro 100  
Db 306 ATCTTGAGCTTCACTCACTAAATGGGGTAAATAGTGTGCTCAAAATTTGCAACAGTTCGA 365  
Qy 101 GlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIleGlnLysHisLys 120  
Db 366 GGAAGAACGATGACGATCAAGAACTATTTGAGAGAAAGAGTTCAAAACAGCTCTGT 425  
Qy 121 GlnAlaGluAsnPhenGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140  
Db 426 CAACCTC---AACAATGAA-----TCTAACAGGCAACAAGTTCTTTGAC---GCTGTT 470  
Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProCysTyr 160

```
Db 471 CGTACTTTTGGTCCCTAGATTGATCGAAGATGGAACAACTCA----- 518
Qy 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
Db 519 -----TTCACCTACTACTACTAT 536
Qy 181 CysCysThrAsnAspAsnAsnIleAsnTyTrp-----SerMetGluAspSer 197
Db 537 TGTGTCTCCCAACAAACAACAACAACACTCTCTCTTCTTCTCTTCATCTCAGACTCT 596
Qy 198 TrpSerMetGln 201
Db 597 TTAAGTATGCAA 608

RESULT 9
ABK65252
ID ABK65252 standard; cDNA; 961 BP.
XX
AC ABK65252;
XX
DT 02-JUL-2002 (first entry)
XX
DE Arabidopsis cDNA encoding a transcription factor #104.
XX
KW Plant; ss; gene; transcription factor; transgenic;
KW agriculture; metabolic chemical; environmental stress; drought;
KW microbial disease resistance; herbicide resistance; seed yield;
KW fruit yield; growth rate; leaf senescence; flower senescence.
XX
OS Arabidopsis thaliana.
XX
PN W020215675-A1.
XX
PD 28-FEB-2002.
XX
PF 22-AUG-2001; 2001WO-US26189.
XX
PR 22-AUG-2000; 2000US-227439P.
PR 16-NOV-2000; 2000US-0713994.
PR 16-APR-2001; 2001US-0837944.
XX
PA (MENDEL) MENDEL BIOTECHNOLOGY INC.
PA (PIG/) PILGRIM M.
PA (CREE/) CREELMAN R.
PA (DUBE/) DUBELL A J.
PA (HEAR/) HEARD J.
PA (JIAN/) JIANG C.
PA (KEDD/) KEDDIE J.
PA (ADAM/) ADAM L.
PA (RATC/) RATCLIFF O.
PA (REUB/) REUBER J L.
PA (RIEC/) RIECHMANN J L.
PA (YUGG/) YU G.
PA (PINE/) PINEDA O.
XX
PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J,
PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;
XX
DR WPI; 2302-292022/33.
XX
DR P-PSDB; AAU93066.
XX
PT An isolated or recombinant polynucleotide used to produce a transgenic
PT plant.
XX
PS Claim 4; Page 443-445; 941pp; English.
XX
CC The invention relates to 1 of 232 isolated or recombinant polynucleotides
CC encoding an Arabidopsis thaliana transcription factor, their variants,
CC complements, fragments, or related polynucleotide with 31% to 95%
CC sequence identity, where the plant possesses an altered trait as compared
CC to a wild-type or reference plant, or the plant exhibits an altered
CC phenotype as compared to a wild-type or reference plant, or the plant
```

```
CC exhibits ectopic expression or altered expression of one or more genes
CC associated with a plant trait as compared to a wild plant. Also included
CC are a transgenic plant comprising the polynucleotides, a computer
CC readable medium having stored sequence information, and identifying a
CC homologue sequence from a database comprising a plurality of known plant
CC sequences comprising inputting sequence information selected from one of
CC 464 fully defined sequences given in the specification. The isolated or
CC recombinant polynucleotide is used for producing a plant having a
CC modified trait, the method comprising selecting a polynucleotide that
CC encodes a polypeptide or an antisense nucleic acid, inserting the
CC polynucleotide or an antisense nucleic acid into an expression vector.
CC Introducing the vector into a plant or a cell of a plant to overexpress
CC the polypeptide or antisense nucleic acid, thereby producing a modified
CC plant, and selecting for a modified trait (e.g. increased
CC production of agriculturally useful proteins or metabolic chemicals,
CC pest tolerance, environmental stress response (e.g. drought), microbial
CC disease resistance, herbicide resistance, seed and fruit yield, growth
CC rate, leaf and flower senescence and many other traits listed in the
CC specification). The present sequence is one of the 232 polynucleotides
CC encoding an A. thaliana transcription factor.
XX
```

Sequence 961 BP; 319 A; 193 C; 235 G; 234 T; 0 other;

## Alignment Scores:

Pred. No.:	2,81e-48	Length:	981
Score:	523.50	Matches:	119
Percent Similarity:	55.77%	Conservative:	26
Best Local Similarity:	45.77%	Mismatches:	32
Query Match:	46.00%	Indels:	83
DB:	24	Gaps:	9

US-10-021-811-36 (1-206) x ABK65252 (1-981)

```
Qy 8 LysThrSerGluAspProGluValArgLysGlyProTPrThrMetGluGluAspLeuIle 27
Db 90 AAAGTGAAGAAAGAAATGACCTTAAGAGAGAGGCTCCATGACAGCTTGAAGAAATTTAGAG 149
Qy 28 LeuMetAsnTyTrpIleAlaAsnHisGlyGluValTTPAsnSerLeuAlaAla 47
Db 150 CTCATCAATTACATGCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 209
Qy 48 GlyLeuLysArgAsnGlyLysSerCysArgLeuArgTPrLeuAsnTyTrpAspProAsp 67
Db 210 GAATCTCAAAAGACCGGAAAGAGCTGACGACTTGGTGGCTGAATATCTCGACACGAT 269
Qy 68 ValArgArgGlyAsnIleThrProGluGluGluLeuIleMetGluLeuHisAlaLys 87
Db 270 GTGCCCGCTGGAACATTAACCTCGAAGAACAACTTGAATTTGAACTTCACACACGCT 329
Qy 88 TrpGlyAsnArgTrpSerLysIleAlaLysHisLeuProGlyArgTrpAspAsnGluIle 107
Db 330 TGGGGCAATAGATGCTTAAGTTGCAACATATTTACACAGAGAAACGATTAAGACGATC 389
Qy 108 LysAsnTyTrpArgTrpArgLleGlnLysHisIleLysGlnAlaGlu-----Asn 124
Db 390 AAAAAGCTATTTGGAACACGTTCTCAAAAGCATGCAAAAAGCTTAAATGCGACGTGAAC 449
Qy 125 PheGlnGln----- 127
Db 450 AGTCAACAATTTAAAGACACCATGAAGATCTTGAATCCTCGGCTCGTAGAAGATC 509
Qy 128 -----GlnSerSerAsn 132
Db 510 CAAGCCGCTCCATCGGCTGTGTTCCATGTCATCTTGGCTGACACCTCTCGATCAG 569
Qy 133 SerGluIleAsnAspHisGlnAlaSerThrSerHisValSerThrMetAlaGluProMet 152
Db 570 TTCGTGATCAACAACAAC-----AACACCAACAACGGATTAATTGGCT----- 614
Qy 153 GluMetTyTrpSerProProCysTyGlnGlyMetLeuGluProPheSerThrGlnPhePro 172
Db 615 TTAATGATTAACCTTAATGTTAC----- 638
```



## RESULT 11

AAH87724  
ID AAH87724 standard; cDNA; 626 BP.  
XX  
AC AAH87724;  
XX  
DT 25-SEP-2001 (first entry)  
XX  
DE Peppermint plant oil gland expressed cDNA 80.  
XX  
KW Peppermint, plant oil gland cell; terpenoid essential oil; resin;  
KM genetic mapping; antisense suppression; recombinant expression; ss.  
XX  
OS Mentha x piperita.  
XX  
PN WO2001.33319-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 19-JAN-2001; 2001WO-US02567.  
XX  
PR 20-JAN-2000; 2000US-0177264.  
XX  
PS (CROT/) CROTEAU R. B.  
PA (LANG/) LANGE B. M.  
PA (WILD/) WILDUNG M. R.  
XX  
PI Croteau RB, Lange BM, Wildung MR;  
XX  
DR WPI; 2001-488706/53.  
XX  
PT New nucleic acid molecules corresponding to mRNA molecules expressed in  
PT terpenoid oil glands for enhancing expression of plant oil gland cell  
PT proteins -  
XX  
PS Claim 1; Page 108; 25pp; English.  
XX  
CC The invention relates to nucleic acid molecules (AAH87645-AAH88116) that  
CC correspond to all or part of a mRNA molecule expressed in plant oil  
CC gland cells, especially peppermint and plant oil glands that produce  
CC terpenoid essential oils and resins. The nucleic acids are useful for  
CC genetically mapping a plant genome for genes expressed in plant oil  
CC gland cells and to suppress (for example by antisense suppression) or  
CC enhance their expression (for example by genetically transforming a  
CC plant cell with a replicable expression vector that expresses one or more  
CC proteins naturally expressed in plant oil gland cells). The nucleic acids  
CC are also useful for recombinant expression of plant oil gland proteins  
CC required for terpenoid essential oil and/or resin production in bacterial  
CC and/or yeast cells.  
XX  
SQ Sequence 626 BP; 187 A; 158 C; 163 G; 118 T; 0 other;

## Alignment Scores:

Pred. No.: 2.16e-45 Length: 626  
Score: 495.00 Matches: 98  
Percent Similarity: 69.59% Conservative: 21  
Best Local Similarity: 57.31% Mismatches: 23  
Query Match: 43.50% Indels: 29  
DB: 22 Gaps: 4

US-10-021-811-36 (1-206) x AAH87724 (1-626)

QY 6 GlnCysLysThrSerGlnAspPro-----GluValArgLysGlyProTyr 20  
Db 116 AGGTGTGAGAGAGAGAGAGAGAGAAATATGATGAGGTGAGAGAGAGGCGCGTGG 175  
QY 21 ThrMetGluGluAspLeuIleleuMetArgTyrIleAlaAsnHISGluGluGlyValTyr 40  
Db 176 AATAGTTGAGAGAGACTTATCACTATCACTACATCGCCCATCAGCGGAGAGCGCATGG 235  
QY 41 AansrLeuAlaLysAlaAlaGlyLeuLysArgGlnGlyLysSerCysArgLeuArgTyr 60

Db 236 AACTCTGTGCTGCTGCACAGAGCCCTCAACAGAACTCGAAGAAAGCTGCAGACTGAGATCG 295  
QY 61 LeuAsnTyrLeuArgProAspValArgArgGluAsnIleThrProGluGluGluLeu 80  
Db 296 CTGAACCTATCTCGCCCGCATGCGACCGTGCAGACATCACTCTTGAAGAGAGAGCTTTTG 355  
QY 81 IleMetGluLeuHISAlaLysTyrGluAsnArgTyrPsrIysIleAlaLysHISLeuPro 100  
Db 356 ATTCTGACCTCCATCTCTCGATCGGCGCAACAGTGTGCAAAATGCGCAGATCTGCCG 415  
QY 101 GLYArgThrAspAsnGluIleLysAsnTyrTyrArgThrArgIleGluLysHISLeuLys 120  
Db 416 GGAAGAAGACGACTACGAACTAAAGAACTACGAGAAACAGAGTGCAA----- 463  
QY 121 GlnAlaGluAsnPhenGlnGlnSerSerAsnAsnSerGluIleAsnAspHISGlnAla 140  
Db 464 -----AACATGCAAGACAGCTCAAT-----GTACGTCACAGACAGCA 502  
QY 141 -SerThrSerHisVal-----SerThrMe 148  
Db 503 TTCAGAGACCCATGCCCTTACCTTTGAGGCTTACGCTGTTGAGAAATCCAAAGCAGCTTC 562  
QY 148 talagluPromeGluMetLysSerProPro 158  
Db 563 TGCATCTGCTGCTCGCTCCGCGCTCCGCTCCG 593  
RESULT 12  
ID AAC57194 standard; DNA; 524 BP.  
XX  
AC AAC57194;  
XX  
DT 25-JAN-2001 (first entry)  
XX  
DE Eucalyptus grandis transcription factor DNA sequence #700.  
XX  
KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
KW basic helix-loop-helix zipper; homeotic; homeobox; MADS;  
KW homeodomain zipper; LIM domain; AP2; ERBs; zinc finger domain;  
KW type 2 Cys2His2; CCAAT box element; MYB; ss.  
XX  
OS Eucalyptus grandis.  
XX  
PN WO200053724-A2.  
XX  
PD 14-SEP-2000.  
XX  
PF 09-MAR-2000; 2000WO-US06112.  
XX  
PR 11-MAR-1999; 99US-0266513.  
PR 18-AUG-1999; 99US-0149485.  
XX  
PA (GENE-) GENESTS RES & DEV CORP LTD.  
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
XX  
PI Wood M, McGrath A, Shenk MA, Glenn M;  
XX  
DR WPI; 2000-579369/54.  
XX  
PT New isolated polynucleotide encoding a plant transcription factor for  
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
PT having modified gene expression or modified activity of a polypeptide  
PT -  
XX  
PS Claim 1; Pages 614-615; 747bp; English.  
XX  
CC The present invention relates to novel plant transcription factors from  
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
CC sequence for one such transcription factor. The transcription factor may  
CC be used to produce a plant having modified gene expression such as a  
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
CC mahogany species or to modify the activity of a polypeptide in a plant.



CC The transcription factors of the present invention are members from the  
CC following families of regulatory proteins: bZIP, bZIP family of G-box  
CC binding factors, basic helix-loop-helix zipper, LIM domain, AP2  
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
CC and MYB.

XX Sequence 524 BP; 143 A; 131 C; 137 G; 113 T; 0 other;

## Alignment Scores:

Pred. No.:	1,29e-43	Length:	524
Score:	478.00	Matches:	81
Percent Similarity:	84.87%	Conservative:	20
Best Local Similarity:	68.07%	Mismatches:	18
Query Match:	42.00%	Indels:	0
DB:	21	Gaps:	0

US-10-021-811-36 (1-206) x AAC57194 (1-524)

QY 3 Lyslysginglncylserthserglnaspprogluvalarglysglyprotpthmet 22  
DB 154 AGGAAGATGACATCTCCGGAGAGAGAGGGGTGACCTCGGAGGGGGCCATGACTCC 213  
QY 23 Glugluaspleulleleumetasnlyrllaalasnhsiglygluglyvalttrpaser 42  
DB 214 GAGGAAGCAATTGTGCTCATCTCCGATCATGCGCCAGGGGCGCTGGAATATG 273  
QY 43 leuialysalaaaglyleuileysasnlysserasyargleuargtrpleuasn 62  
DB 274 TTGGGAAAGAGCGGAGATTGAAGAGAACTGGCAAAACCTGAGTTAAGGTGGTGAAT 333  
QY 63 Tyrleuargproaspvalargargglyasnillethrpogluglulinleuileu 82  
DB 334 TACCTGAGACCCGAGATCAAGGGCGGAGATCTCACCCGCAAGAACACTCATGATCCTT 393  
QY 83 Gluleuhsalalyserlyglsasnargtrpserlylleaalyshisleuarg 102  
DB 394 GAACCTTCAACCAATGGGGCAACAGGTGGTCGAAATCGGGCACTATCTCCCGAGAGG 453  
QY 103 Thrappangluillelysasnlyrtrpargtrprrargllleuileys 121  
DB 454 ACGATTAACGAGATCAAGAACTACTGAGAGACGGGGTGCAGAACCAAGCGCGCAT 510

## RESULT 13

AAX25572 standard; CDNA; 837 BP.

AA25572;

02-AUG-1999 (first entry)

Arabidopsis thaliana MYB2 (AtMYB2) cDNA.

MYB2 gene; AtMYB2 gene; transcription factor; transgenic plant;  
KM environmental stress; stress tolerance; anaerobic stress;  
KM drought; flooding; salt; cold; crop protection; ss.

OS Arabidopsis thaliana.

Key Location/Qualifiers  
FT CDS 10..831  
FT /tag= a

MO9916878-A1.

08-APR-1999.

25-SEP-1998; 98WO-AU00812.

26-SEP-1997; 97AU-0009479.

PA (CSTR ) COMMONWEALTH SCI & IND RES ORG.

PI Dennis ES, Dolferrus RAM, Hoeren FU, Peacock WJ;

XX WPI; 1999-263695/22.

DR P-PSDB; AAY05831.

PT Altering expression of a stress-related gene in a plant

XX Claim 22; Page 72-74; 119pp; English.

CC This is the sequence of Arabidopsis thaliana ecotype C24 cDNA  
CC encoding the stress-related protein AtMYB2 (see AAY05831). The  
CC cDNA was obtained by RT-PCR of anaerobically induced root RNA  
CC derived from ecotype C24. The inventors have demonstrated that a  
CC family of transcription factors, the MYB2 family of proteins, are  
CC capable of inducing the expression of stress-related genes in plant  
CC cells in response to stress situations. The MYB2 protein encoded  
CC by the AtMYB2 gene is useful for the purposes of inducing or  
CC repressing the expression of plant-expressible genes that are  
CC involved in the plant's response to anaerobic stress, flooding  
CC stress, cold stress, dehydration stress, drought stress, heat  
CC stress or salinity. The stress-related gene, such as the MYB2  
CC alcohol dehydrogenase Adh1 gene, is transactivated by the MYB2  
CC polypeptide by virtue of the presence of at least one copy of a  
CC cis-acting regulatory sequence, in particular a MBS-1 (Myb binding  
CC site) motif and/or a GC-motif and/or a G-box-1 motif and/or a  
CC G-box-2 motif, as found in the Adh1 promoter (see AAX25573-74). The  
CC MYB2 protein is able to target several stress-induced enzymes,  
CC rather than targeting just one, allowing the production of stress  
CC tolerant transgenic plants.

XX Sequence 837 BP; 246 A; 176 C; 212 G; 203 T; 0 other;

## Alignment Scores:

Pred. No.:	5.31e-43	Length:	837
Score:	475.00	Matches:	92
Percent Similarity:	66.48%	Conservative:	27
Best Local Similarity:	51.40%	Mismatches:	36
Query Match:	41.74%	Indels:	24
DB:	20	Gaps:	4

US-10-021-811-36 (1-206) x AAX25572 (1-837)

QY 11 Gluaspprogluvalarglysglyprotpthmerglugluaspleulleuemetasn 30  
DB 55 GAAGATTCTGATGATGAGAAAGGTCCATGAGACGAGAGAAAGATCAATCCATGCAAC 114  
QY 31 Tyrillealsnhsiglygluglyvalttrpaserleuileuileuileuileu 50  
DB 115 TTGCTCTTATTATGAGCGATGCTCGTTGAAACCAATGCTGCTGCGCTAAAG 174  
QY 51 Argasnlylysserasyargleuargtrpleuasnlyrleuargproaspvalargarg 70  
DB 175 CGAAGCTGTAAGAGTTGATTAAGATGGCTTAATTAATTAATTAATTAATTAAGAG 234  
QY 71 Glyasnillethrpogluglulinleuileuileuileuileuileuileuileuileu 90  
DB 235 GGCACATCACTCGAAGAACCAATTATGATCTCAAACTCATCTCTTGGGGCAAT 294  
QY 91 Argtrpserlylleaalyshisleuargtrprrargllleuileuileuileuileu 110  
DB 295 AGGTGTCGAAAGATGGCGCAATATCTACCGGAGAGAGAGATTAATTAATTAATTAAT 354  
QY 111 Trrpargtrprrargllleuileuileuileuileuileuileuileuileuileu 125  
DB 355 TGGAGACTCGAGTCAAGAACCAACCACTTAAGATGAGTGAACAGTAACTT 414  
QY 126 Glnglnglnseraseranapn-----serglulleasnaphis 138  
DB 415 TTCAAGAGACTATGAGAAATGTTTGATGCCAGATTGAGAAAGAAAGAGCCCA 474  
QY 139 Glualaserthr-----serhivalserthmetaglupromet 152  
DB 475 TCATTACCCACGTCGTGAACAAGTGAGTCAATGATACGACCAACCAAGTCAACAGTT 534





```

XX (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
XX Wood M, McGrath A, Shenk MA, Glenn M,
PI WPI: 7000-579369/54.
XX
XX New isolated polynucleotide encoding a plant transcription factor for
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT having modified gene expression or modified activity of a polypeptide
PT
XX
XX Claim 1; Page 120; 747pp; English.
XX
XX The present invention relates to novel plant transcription factors from
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding
CC sequence for one such transcription factor. The transcription factor may
CC be used to produce a plant having modified gene expression such as a
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or
CC mahogany species or to modify the activity of a polypeptide in a plant.
CC The transcription factors of the present invention are members from the
CC following families of regulatory proteins: bzip, bzip family of G-box
CC binding factors, basic helix-loop-helix zipper, LIM domain, AP2
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2
CC and ERBBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements
CC and MYB.
XX
SQ Sequence 389 BP; 112 A; 102 C; 97 G; 78 T; 0 other;

Alignment Scores:
Pred. No: 1,75e-37 Length: 389
Score: 421.00 Matches: 81
Percent Similarity: 88.17% Conservative: 1
Best Local Similarity: 87.10% Mismatches: 9
Query Match: 36.99% Indels: 2
DB: 21 Gaps: 1

US-10-021-811-36 (1-206) x AAC56152 (1-389)

QY 1 MetAspIysLys-----GInGInCysLysThrSerGInAspProGInuValArgLysGly 18
Db 44 ATGGACAAAGAGCCAGACGACGAGTGTAGTCCCAAGATGTCGAGGTGAGAAAAAGG 103
QY 19 ProTrpThrmMetGInuAspLeuIleLeuMetAsnTrpIleAlaAsnHisGlyGInGly 38
Db 104 CCGTGGACGATGAGAGGATCTCATCTCATCACTAAGGATCAGGCGAAGGC 163
QY 39 ValTrpAsnSerLeuAlaLysAlaIleGlyLeuLysArgAsnGlyLysSerCysArgLeu 58
Db 164 AGTTGAACTCCCTAGCCAAAGCTGCTGTCTTAAACGTACCGGAAAGATTGTCGGCTC 223
QY 59 ArgTrpLeuAsnTrpLeuArgProAspValArgArgGlyAsnIleThrProGInuGInGIn 78
Db 224 CCGTGGCTGAACATATCTGCGACCCGACGTCGAGAGGCAACATCACTACTGAGAGCAG 283
QY 79 LeuLeuIleMetGInuLeuHisAlaLysTrpGlyAsnArg 91
Db 284 CTCTGATCATGGAACATGATGCCAAGTGGGAAACAGG 322

```

Search completed: February 8, 2003, 11:12:19  
Job time : 226 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: February 8, 2003, 11:07:40 ; Search time 48 Seconds

(without alignments)  
1316.154 Million cell updates/sec

Title: US-10-021-811-36

Perfect score: 1138

Sequence: 1 MDKQCKTSQDPYVKKGPW.....NNINYSMEDSWSMOJLNGD 206

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DRV=x1h  
-Q/cgn2\_1/USFTO.spool/US10021811/runat.03022003.111222.25655/app.query.fastc\_1.391  
-DB=Issued\_Patents\_NA -OPMT=fastc -SOFTX=rni -WIMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOCMATCH=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10021811 @CGN 1.1 25 @runat.03022003.111222.25655 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAR -LARGEJOBRY -NEG\_SCORES=0 -WAIT -LONGLOG -DRV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database :

1: Issued Patents NA:\*  
2: /cgn2\_6/ptodata/2/ina/5A COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/5B COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6A COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/6B COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	396	34.8	2220	4 US-08-997-251-1	Sequence 1, Appl1
2	386	33.9	2352	4 US-08-997-251-3	Sequence 3, Appl1
3	382.5	33.6	1344	2 US-08-732-6268-1	Sequence 1, Appl1
4	305.5	26.8	1035	5 PCT-US93-06251-90	Sequence 90, Appl1
5	305.5	26.8	3225	1 US-08-306-6918-45	Sequence 45, Appl1
6	305.5	26.8	3225	5 PCT-US93-06251-91	Sequence 91, Appl1
7	305.5	26.8	3230	5 PCT-US93-06251-89	Sequence 89, Appl1
8	296	26.0	1897	4 US-09-167-322-10	Sequence 10, Appl1
9	293	25.7	4880	4 US-09-402-929-5	Sequence 5, Appl1
10	291.5	25.6	3602	4 US-09-402-929-1	Sequence 1, Appl1
11	291.5	25.6	6775	4 US-09-402-929-4	Sequence 4, Appl1
12	271.5	23.9	2638	1 US-08-306-6918-46	Sequence 46, Appl1

13	248	21.8	4059	2 US-08-485-139-1	Sequence 1, Appl1
14	248	21.8	4059	3 US-08-750-357-1	Sequence 1, Appl1
15	248	21.8	4824	2 US-08-485-139-5	Sequence 5, Appl1
16	248	21.8	4824	3 US-08-750-357-5	Sequence 5, Appl1
17	185.5	16.3	469	3 US-09-008-979A-2	Sequence 2, Appl1
18	185.5	16.3	469	4 US-09-460-618-2	Sequence 2, Appl1
19	185.5	16.3	469	4 US-09-310-235B-2	Sequence 2, Appl1
20	165.5	14.5	5889	4 US-09-402-929-3	Sequence 3, Appl1
21	125	11.0	785	3 US-09-008-979A-6	Sequence 6, Appl1
22	125	11.0	785	4 US-09-460-618-6	Sequence 6, Appl1
23	125	11.0	785	4 US-09-310-235B-6	Sequence 6, Appl1
24	119	10.5	2837	4 US-09-156-316-11	Sequence 11, Appl1
25	105.5	9.3	1116	4 US-08-928-941D-17	Sequence 17, Appl1
26	105.5	9.3	1116	4 US-09-280-590A-17	Sequence 17, Appl1
27	105.5	9.3	2903	4 US-08-928-941D-2	Sequence 2, Appl1
28	105.5	9.3	2903	4 US-08-928-941D-3	Sequence 3, Appl1
29	105.5	9.3	2903	4 US-09-280-590A-2	Sequence 2, Appl1
30	105.5	9.3	2903	4 US-09-280-590A-3	Sequence 3, Appl1
31	105.5	9.3	3767	4 US-08-928-941D-28	Sequence 28, Appl1
32	105.5	9.3	3767	4 US-08-928-941D-30	Sequence 30, Appl1
33	105.5	9.3	3767	4 US-09-280-590A-28	Sequence 28, Appl1
34	105.5	9.3	3767	4 US-09-280-590A-30	Sequence 30, Appl1
35	96.5	8.5	282	2 US-08-814-030-2	Sequence 2, Appl1
36	93	8.2	336	1 US-08-248-474-105	Sequence 105, App
37	93	8.2	336	3 US-08-756-849-105	Sequence 105, App
38	84.5	7.4	3834	4 US-09-221-017B-663	Sequence 663, Appl
39	83.5	7.3	2581	4 US-09-370-838-66	Sequence 66, Appl
40	82.5	7.2	3642	3 US-08-946-026-16	Sequence 16, Appl
41	81.5	7.2	3627	1 US-08-447-500-23	Sequence 23, Appl
42	81.5	7.2	3627	1 US-08-454-037-23	Sequence 23, Appl
43	81.5	7.2	3627	1 US-08-453-866-23	Sequence 23, Appl
44	81.5	7.2	3627	3 US-08-185-359-23	Sequence 23, Appl
45	81.5	7.2	3628	1 US-08-468-036-17	Sequence 17, Appl

#### ALIGNMENTS

RESULT 1  
US-08-997-251-1  
; Sequence 1, Application US/08997251  
; Parent No. 6271440  
; GENERAL INFORMATION:  
; APPLICANT: GUBLER, FRANZ J.  
; APPLICANT: JACOBSEN, JOHN V.  
; TITLE OF INVENTION: PLANT REGULATORY PROTEINS III  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Boulevard  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/997,251  
; FILING DATE: 23-DEC-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO AU96/00383  
; FILING DATE: 21-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PN6470/95  
; FILING DATE: 09-NOV-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PN3779/95  
; FILING DATE: 23-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winner, Ellen F.

```

/ / REGISTRATION NUMBER: 28,547
/ / REFERENCE/DOCKET NUMBER: 110-97
/ / TELECOMMUNICATION INFORMATION:
/ / TELEPHONE: (303) 499-8080
/ / TELEFAX: (303) 499-8089
/ / INFORMATION FOR SEQ ID NO: 1:
/ / SEQUENCE CHARACTERISTICS:
/ / LENGTH: 2220 base pairs
/ / TYPE: nucleic acid
/ / STRANDEDNESS: single
/ / TOPOLOGY: linear
/ / MOLECULE TYPE: cDNA
/ / ORIGINAL SOURCE:
/ / ORGANISM: Hordeum vulgare
/ / STRAIN: Himalaya
/ / DEVELOPMENTAL STAGE: Seed
/ / TISSUE TYPE: Aleurone layer
/ / IMMEDIATE SOURCE:
/ / CLONE: HvGAMYb
/ / FEATURE:
/ / NAME/KEY: CDS
/ / LOCATION: 275..1933
/ /
US-08-997-251-1
/ /
Alignment Scores:
Pred. No.: 1.47e-38 Length: 2220
Score: 396.00 Matches: 70
Percent Similarity: 69.50% Conservative: 28
Best Local Similarity: 49.65% Mismatches: 31
Query Match: 34.80% Indels: 12
DB: 4 Gaps: 1

US-10-021-811-36 (1-206) x US-08-997-251-1 (1-2220)
QY 13 ProgluValArgLyGlyProTrpThrMetGluGluAspLeuIleLeuMetAsnTyrIle 32
Db 392 CGGCTCTGAAGAGAGGGGCGCTGACGTCGCCGAGAGACGCCATCTCTGGAGCTACAGTG 451
QY 33 AlaAenHisGlyGluGlyValTrpAsnSerIleuAlaIleGlyLeuIleYAsnArgAsn 52
Db 452 AAGAAACACGGGAGAGGAGAACTGGAACCGGGTGCAAGAAACACCGGGCTTTCGCCGTGC 511
QY 53 GlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArgGlyAsn 72
Db 512 GGCAGAGAGCTGGCGCTCCGGTGGCGAACCACCTCAGGCCCACTCAAGAAAGGGGCC 571
QY 73 ILeTrpProGluGluGluLeuIleMetGluLeuHisAlaIleYTrpGlyAsnArgTrp 92
Db 572 TTCACCCCGAGAGAGAGAGGCTCATCACTCCAGCTCCCAAGATGGGCAACAAGTGC 631
QY 93 SerLysIleAlaIleYHisIleuProGlyArgThrAspAsnGluIleYAsnTyrTrpArg 112
Db 632 GCTCGATGGCGGCTCATTTGGCAGGGGTACTGATTAATGAATTAAGAAATTAAGTCTGAGAC 691
QY 113 ThrArgGluGlnLys-----HisIleLys 120
Db 692 ACTCGAATTAAGAGATGTCACGAGACCGGTTTGCCAAATATATCTCTGATGTATGCAAT 751
QY 121 GlnAlaGluAspHseGluGlnIleuSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db 752 CAATCTTCAATGAAGATCAGAGAGGCTCCAGCATTTCAACTGGGGGGAATCTTTCC 811
QY 141 Ser 141
Db 812 AsT 814

```

```

/ / TITLE OF INVENTION: PLANT REGULATORY PROTEINS III
/ / NUMBER OF SEQUENCES: 17
/ / CORRESPONDENCE ADDRESS:
/ / ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
/ / STREET: 5370 Manhattan Boulevard
/ / City: Boulder
/ / STATE: CO
/ / COUNTRY: US
/ / ZIP: 80303
/ / COMPUTER READABLE FORM:
/ / MEDIUM TYPE: Floppy disk
/ / COMPUTER: IBM PC compatible
/ / OPERATING SYSTEM: PC-DOS/MS-DOS
/ / SOFTWARE: Patent Release #1.0, Version #1.30
/ / CURRENT APPLICATION DATA:
/ / APPLICATION NUMBER: US/08/997,251
/ / FILING DATE: 23-DEC-1997
/ / CLASSIFICATION: 435
/ / PRIOR APPLICATION DATA:
/ / APPLICATION NUMBER: WO A/96/00383
/ / FILING DATE: 21-JUN-1996
/ / PRIOR APPLICATION DATA:
/ / APPLICATION NUMBER: AU P/6470/95
/ / FILING DATE: 09-NOV-1995
/ / PRIOR APPLICATION DATA:
/ / APPLICATION NUMBER: AU P/3779/95
/ / FILING DATE: 23-JUN-1995
/ / ATTORNEY/AGENT INFORMATION:
/ / NAME: Winner, Ellen P.
/ / REGISTRATION NUMBER: 28,547
/ / REFERENCE/DOCKET NUMBER: 110-97
/ / TELECOMMUNICATION INFORMATION:
/ / TELEPHONE: (303) 499-8080
/ / TELEFAX: (303) 499-8089
/ / INFORMATION FOR SEQ ID NO: 3:
/ / SEQUENCE CHARACTERISTICS:
/ / LENGTH: 2352 base pairs
/ / TYPE: nucleic acid
/ / STRANDEDNESS: single
/ / TOPOLOGY: linear
/ / MOLECULE TYPE: cDNA
/ / ORIGINAL SOURCE:
/ / ORGANISM: Oryza sativa
/ / FEATURE:
/ / NAME/KEY: CDS
/ / LOCATION: 396..2054
/ /
US-08-997-251-3
/ /
Alignment Scores:
Pred. No.: 2.69e-37 Length: 2352
Score: 386.00 Matches: 68
Percent Similarity: 69.57% Conservative: 28
Best Local Similarity: 49.28% Mismatches: 30
Query Match: 33.92% Indels: 12
DB: 4 Gaps: 1

US-10-021-811-36 (1-206) x US-08-997-251-3 (1-2352)
QY 13 ProgluValArgLyGlyProTrpThrMetGluGluAspLeuIleLeuMetAsnTyrIle 32
Db 507 CCCCCCTGAAGAGAGGGGCAATGACGTCGGGAGAGAGCCCATCTGATGACTACAGTG 566
QY 33 AlaAenHisGlyGluGlyValTrpAsnSerIleuAlaIleGlyLeuIleYAsnArgAsn 52
Db 567 AAGAGACACGGGAGAGGGAACGTGAACCGGGTGCAAGAAACACCGGGCTGTCGGGTGC 626
QY 53 GlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArgGlyAsn 72
Db 627 GGCAGAGCTTCCGCTCCGGTGGCGAACCACCTGAGGCCCACTCAAGAAAGGGGCC 686
QY 73 ILeTrpProGluGluGluLeuIleMetGluLeuHisAlaIleYTrpGlyAsnArgTrp 92
Db 687 TTCACCGCGAGAGAGAGGCTCATCACTCCAGCTCCCAAGATGGGCAACAAGTGC 746

```

Qy 93 SerIleAlaIysHisLeuProGlyArgThrAspAsnGluIleIysAsnTyrTrpArg 112  
Db 747 GCTGGATGCGCCCTCTATTGCGAGCGCATGATATGAATTAAGATTCTGGAA 806  
Qy 113 ThrArgIleGlnIys-----HisIleIys 120  
Db 807 ACTCGAATTAAGAGATCCAGCGAGCTGACCTATCTATCTATCCAGCGTATGCAAT 866  
Qy 121 GlnAlaGluLeuPheGlnGlnIleSerSerAsnAsnSerGluIleAsnAspHis 138  
Db 867 CAATCTCAATTAAGAGATGAGATGCTGCTCAATGATTTGACTGTGGCGAAT 920  
RESULT 3  
US-08-722-626B-1  
Sequence 1, Application US/08722626B  
Patent No. 593601  
GENERAL INFORMATION:  
APPLICANT: Yang, Yining  
APPLICANT: Kleesig, Daniel, F.  
TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ENHANCED  
DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dann, Dorfman, Herrell and Skillman  
STREET: 1601 Market Street Suite 720  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2307  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/722,626B  
FILING DATE: 27-SEP-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pat Hagan  
REGISTRATION NUMBER: 27,643  
REFERENCE/DOCKET NUMBER: 97-0010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215 563-4100  
TELEFAX: 215 563-4044  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1344 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 148...981  
OTHER INFORMATION:  
US-08-722-626B-1  
Alignment Scores:  
Pred. No.: 3.09e-37 Length: 1344  
Score: 382.50 Matches: 88  
Percent Similarity: 54.38% Conservative: 30  
Best Local Similarity: 40.55% Mismatches: 60  
Query Match: 33.61% Indels: 39

DB: 2 Gaps: 6  
US-10-021-811-36 (1-206) x US-08-722-626B-1 (1-1344)  
Qy 15 ValArgIysGlyProTyrThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsn 34  
Db 181 CTGAAGAAAGGCGCATGGATTCCTGGAAGATGATGATTCATCTCTTTCATTCAAACT 240  
Qy 35 HisGlyGluGlyValTyrAsnSerIleuAlaIleGlyLeuIleAsnGlyIys 54  
Db 241 AATGCGCATGGCACTGGCGAGCCCTTCCAAACAGCGTGAGTATGAGATGCGGGAAG 300  
Qy 55 SerCyAsnGluAsnTyrIleuAsnTyrIleuArgProAspValArgArgIleAsnIleThr 74  
Db 301 AGTTGACAGCTGCGGTGAGCGAATTTTGGACCAATATTAAGAGAGGGAATTTTCC 360  
Qy 75 ProGluGluGlnLeuLeuIleuGluLeuHisAlaIysTyrGlyAsnArgTyrSerIys 94  
Db 361 AAGGAAGAAAGAAACAAATTTCAGTTACATGAATGCTTGGCAATGATGCTGCA 420  
Qy 95 IleAlaIysHisLeuProGlyArgThrAspAsnGluIleIysAsnTyrTrpArgThrArg 114  
Db 421 ATAGCAGCAAAATTCACGAGCAAGCAATGAAATTAATAATGTTGGCACACCCAC 480  
Qy 115 IleGlnIysHisIleIys-----GlnAla 122  
Db 481 TTGAAGAAAGAACTCAAGATTATTAAGCTCTCCAGAACTCCAAAGACACTCCAAAGTCC 540  
Qy 123 GluAsnPheGln-----GlnIleSerSerAsnAsnSerGluIle--- 135  
Db 541 AAGATCATGATTCAGAGGCTCTACTACTTCTGATCATCCAAATTTGATGATTTTCT 600  
Qy 136 -----AsnAspHisGlnAla 140  
Db 601 ATTATTAATACAAACAAACATGATAGCCAGTGTAGTCTCTTAACCTCAACCCCAATT 660  
Qy 141 SerThrSer---HisValSerThrMetAlaGluProMetGluMetTyrSerProProCys 159  
Db 661 TCATCTTAGTACTGAATGCACTGACACACAGTCAGATGATCATCAATAGTGTGATT 720  
Qy 160 TyrGlnGlyMetLeuGluIleProPheSerThrGlnPheProThrIleAsnProAspGlnSer 179  
Db 721 AAGCAAGAAAGTAATGAGTGC---TCCGAGTATTTTCCAGATGATGAT-----GAG 768  
Qy 180 SerCyAsnGlyThrAsnAspAsnAsnIleAsnTyrTrpSerMetGluAsp 196  
Db 769 AGTTTGGACGAGCAATTAACAACGACAAATTAAGTGAAGTATGAT 819  
RESULT 4  
PCT-US93-06251-90  
Sequence 90, Application PC/TUS9306251  
GENERAL INFORMATION:  
APPLICANT: Wickstrom, Eric and Rife, Jason P.  
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing  
Stereospecific Alkylphosphonates and Arylphosphonates  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: NY  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/06251  
FILING DATE: 19930630  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:

/ NAME: Digiglo, Frank S.  
 / REGISTRATION NUMBER: 31,346  
 / REFERENCE/DOCKET NUMBER: 8586  
 / TELECOMMUNICATION INFORMATION:  
 / TELEPHONE: 516-742-4343  
 / TELEFAX: 516-742-4366  
 / TELEX: 230 901 SANS UR  
 / INFORMATION FOR SEQ ID NO: 90:  
 / SEQUENCE CHARACTERISTICS:  
 / LENGTH: 1035 base pairs  
 / TYPE: nucleic acid  
 / STRANDEDNESS: double  
 / TOPOLOGY: linear  
 / MOLECULE TYPE: DNA (genomic)  
 / PCT-US93-06251-90

Alignment Scores:  
 Pred. No.: 5.58e-28 Length: 1035  
 Score: 305.50 Matches: 70  
 Percent Similarity: 59.34% Conservative: 38  
 Best Local Similarity: 38.46% Mismatches: 59  
 Query Match: 26.85% Indels: 15  
 DB: 5 Gaps: 7

US-10-021-811-36 (1-206) x PCT-US93-06251-90 (1-1035)

```

QY 12 AAPPAGUVALAAGLYGGLYProTPrThMetGluGluAerPLeuileuMetAsnTyr 31
DB 121 AACCTGAGCTCATCAAGGTCCTTGACCAAGAAAGATCAGAGTATAGACTT 180
QY 32 TLealaemniSGLYGUGLYValTPraSerleuAlaLysAlaAGLYLeuLys-- 50
DB 181 CTACAGAAATACGGTCCGAAACGTTGCTGTATGTCAG--CACTTAAAGGG 234
QY 51 ArgAnGlyLysSerCysArgLeuArgTPrLeuAsnTyrLeuA-gProAspValArgArg 70
DB 235 AGAATTGAAAACATGATGAGGAGAGGTCATTAACCACTTGAATCCAGAGTTAAAGAA 294
QY 71 GLYAsnileThrProGluGluGlnLeuLeuileMetGluLeuHISAlaLysTyrGlyAsn 90
DB 295 ACCCTCGAGAGAGAGAGAGAGCAATTTATTCAGGACGACCAAGAGCTGGGGAC 354
QY 91 ArgTPrSerLysIleAlaLysHISLeuProGlyArgTPrAspAnGluileLysAsnTyr 110
DB 355 AGATGGGCAAAATCGCAAGACTAGCTGCTGAGCAAGTATATGCTATCAAGAACAC 414
QY 111 TPrArgThrArgIleGlnLysHISileLysGlnAlaGluAsnPhenGlnGlnSerSer 130
DB 415 TCGAATTCACAAATGCGTCGAGAGTGCAGAG--GAAGTTATTCGACGAGACTTTCA 471
QY 131 AsnAsnSerGlu-----IleAsnAspHisGlnAlaSerThrSerHisValSerThMet 148
DB 472 AAAGCCAGCCAGCCAGAGAGTGCACAAAGCTTCCAGAAAGAAAGTCAATTGATGGGTTT 531
QY 149 ALaGluProMetGluMetTyrSerProProCysTyrGlnGluMetLeuGluProPheSer 168
DB 532 GCTCAG-----GCTCCGCTTACAGCTCA-----CTCCCTGCCACT 567
QY 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAsnAsnAsn 188
DB 568 GGCAG--CCACTGTTAACAAAGACTATCTTATTCACCAATTCCTGAAGCAAAAT 624
QY 189 IleAsn 190
DB 625 GTCTCC 630

```

RESULT 5  
 US-08-306-691B-45  
 ; Sequence 45, Application US/08306691B  
 ; Patent No. 5734039  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Calabretta, Bruno  
 ; APPLICANT: Skorski, Tomasz

/ TITLE OF INVENTION: ANTISENSE  
 / NUMBER OF INVENTION: Oligonucleotides Targeting Cooperating Oncogenes  
 / NUMBER OF SEQUENCES: 55  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P. C.  
 / STREET: Two Penn Center, Suite 1800  
 / CITY: Philadelphia  
 / STATE: Pennsylvania  
 / COUNTRY: U.S.A.  
 / ZIP: 19102  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
 / COMPUTER: IBM PS/2  
 / OPERATING SYSTEM: MS-DOS  
 / SOFTWARE: WordPerfect 5.1  
 / CURRENT APPLICATION DATA:  
 / APPLICATION NUMBER: US/08/306,691B  
 / FILING DATE: September 15, 1994  
 / CLASSIFICATION: 514  
 / PRIOR APPLICATION DATA:  
 / APPLICATION NUMBER:  
 / FILING DATE:  
 / ATTORNEY/AGENT INFORMATION:  
 / NAME: Monaco, Daniel A.  
 / REGISTRATION NUMBER: 30,480  
 / REFERENCE/DOCKET NUMBER: 8321-8  
 / TELECOMMUNICATION INFORMATION:  
 / TELEPHONE: (215) 568-8383  
 / TELEFAX: (215) 568-5549  
 / TELEX: No. 5734039e  
 / INFORMATION FOR SEQ ID NO: 45:  
 / SEQUENCE CHARACTERISTICS:  
 / LENGTH: 3225 base pairs  
 / TYPE: nucleic acid  
 / STRANDEDNESS: double  
 / TOPOLOGY: linear

US-08-306-691B-45

Alignment Scores:  
 Pred. No.: 3.13e-27 Length: 3225  
 Score: 305.50 Matches: 70  
 Percent Similarity: 59.34% Conservative: 38  
 Best Local Similarity: 38.46% Mismatches: 59  
 Query Match: 26.85% Indels: 15  
 DB: 1 Gaps: 7

US-10-021-811-36 (1-206) x US-08-306-691B-45 (1-3225)

```

QY 12 AAPPAGUVALAAGLYGGLYProTPrThMetGluGluAerPLeuileuMetAsnTyr 31
DB 372 AACCTGAGCTCATCAAGGTCCTTGACCAAGAAAGATCAGAGTATAGACTT 431
QY 32 TLealaemniSGLYGUGLYValTPraSerleuAlaLysAlaAGLYLeuLys-- 50
DB 432 GACAGAAATACGGTCCGAAACGTTGCTGTATGTCAG--CACTTAAAGGG 485
QY 51 ArgAnGlyLysSerCysArgLeuArgTPrLeuAsnTyrLeuA-gProAspValArgArg 70
DB 486 AGAATTGAAAACATGATGAGGAGAGTGCATTAACCACTTGAATCCAGAAAGTTAAAGAA 545
QY 71 GLYAsnileThrProGluGluGlnLeuLeuileMetGluLeuHISAlaLysTyrGlyAsn 90
DB 546 ACCCTCGAGAGAGAGAGAGAGCAATTTATTCAGGACGACCAAGAGCTGGGGAC 605
QY 91 ArgTPrSerLysIleAlaLysHISLeuProGlyArgTPrAspAnGluileLysAsnTyr 110
DB 606 AGATGGGCAAAATCGCAAGACTAGCTGCTGAGCAAGTATATGCTATCAAGAAACAC 665
QY 111 TPrArgThrArgIleGlnLysHISileLysGlnAlaGluAsnPhenGlnGlnSerSer 130
DB 666 TCGAATTCACAAATGCGTCGAGAGTGCAGAG--GAAGTTATTCGACGAGACTTTCA 722
QY 131 AsnAsnSerGlu-----IleAsnAspHisGlnAlaSerThrSerHisValSerThMet 148

```



```
Db 723 AAGCCAGCCAGCCAGCAGTGGCCCAAGCTTCCAGAAAGACGTCATTTGATGGGTTT 782
Qy 149 AAGlupPrometGluMetYrserProProCysTyrgIngluMetLeuGluProPheSer 168
Db 783 GCTCAG-----GCTCCGCTTACAGCTCAA-----CTCCCTGCCACT 818
Qy 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsn 188
Db 819 GGCACG---CCCATGTTTACACAGACTATTCCTATTACACATTTCTGAAGACAAAT 875
Qy 189 IleAsn 190
Db 876 GTCCTC 881

RESULT 6
PCT-US93-06251-91
; Sequence 91, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3225 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-91

Alignment Scores:
Pred. No.: 3,13e-27 Length: 3225
Score: 305.50 Matches: 70
Percent Similarity: 59.34% Conservative: 38
Best Local Similarity: 38.46% Mismatches: 59
Query Match: 26.85% Indels: 15
Gaps: 7

US-10-021-811-36 (1-206) x PCT-US93-06251-91 (1-3225)
Qy 12 AspProGluValArgLysGlyProTrpThrMetGluGluAspLeuIleLeuMetAsnTyr 31
Db 372 AACCTGAGCTCATCAGAGGTCCTTGACCAAGAAAGATCAGAGGTGATGAGCTT 431
Qy 32 IleAlaAsnHisGlyGluGluValTrpAsnSerLeuAlaLysAlaLysLys--- 50
Db 432 GTACGAAATACGGTCCGAAACGTTGCTGTATTATGCCAAG-----CACTTAAGGGG 485
```

```
Qy 51 ArgAsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrIleAsnArgProAspValArgArg 70
Db 486 ACAATTGGAAACATGATGAGGAGGTGGCATTAACCACTGAATCCAGAGTTAAGAAA 545
Qy 71 GlyAsnIleThrProGluGluGluLeuLeuIleMetGluMetHisAlaLysTrpGlyAsn 90
Db 546 ACCCTCGTGCAGGAAGAGAGAACAGAAATTATTATCCAGGACACACAGACACTGGGGAAC 605
Qy 91 ArgTrpSerLysIleAlaLysHisLeuProGluArgTrpAspAsnGluIleLysAsnTyr 110
Db 606 AAGTGGCAGAAATCGCAAGACTACTGCTGGACGAATGATTAATCATCAAGAACAC 665
Qy 111 TrpArgThrArgIleGluLysHisIleLysGlnAlaGluAsnGluGlnGlnSerSer 130
Db 666 TCGAATTCTTACAATGCGTCGGAAGGTGCAACAG---GAAGTTATCTGCAGAGCTTCA 722
Qy 131 AsnAsnSerGlu-----IleAsnAspHisGlnAlaSerThrSerHisValSerThrMet 148
Db 723 AAGCCAGCCAGCCAGCAGTGGCCCAAGCTTCCAGAAAGACGTCATTTGATGGGTTT 782
Qy 149 AAGlupPrometGluMetYrserProProCysTyrgIngluMetLeuGluProPheSer 168
Db 783 GCTCAG-----GCTCCGCTTACAGCTCAA-----CTCCCTGCCACT 818
Qy 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsn 188
Db 819 GGCACG---CCCATGTTTACACAGACTATTCCTATTACACATTTCTGAAGACAAAT 875
Qy 189 IleAsn 190
Db 876 GTCCTC 881

RESULT 7
PCT-US93-06251-89
; Sequence 89, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3230 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-89
```

## Alignment Scores:

Pred. No.: 3,14e-27 Length: 3230  
 Score: 305.50 Matches: 70  
 Percent Similarity: 59.34% Conservative: 38  
 Best Local Similarity: 38.46% Mismatches: 59  
 Query Match: 26.85% Indels: 15  
 DB: 5 Gaps: 7

US-10-021-811-36 (1-206) x PCT-US93-06251-89 (1-3230)

QY 12 AspProGluValArgLysGlyProThrPheMetGluGluAspLeuIleLeuMetAsnTyr 31  
 Db 262 AACCTGAGCTCATCAAGGGTCTTGAGCAAGAGAGATCAAGAGATGATAGAGCTT 321  
 QY 32 TleAlaAsnHisGlyGluGlyValTyrAsnSerLeuAlaLysAlaGlyLeuLys--- 50  
 Db 322 GTACAGAAATACGCTCCGAAAGCTTGCTGTATTGCCAAG-----CACTTAAAGGGG 375  
 QY 51 ArgAsnGlyLysSerCysArgLeuArgTyrPheAsnTyrLeuArgProAspValArgArg 70  
 Db 376 AGAATTGGAAAAACAATGAGGAGGTGGCATACCACTTGAAATCCAGAGTTAAGAAA 435  
 QY 71 GlyAsnIleThrProGluGluGlnLeuLeuIleMetGluLeuHisAlaLysTyrGlyAsn 90  
 Db 436 ACCTCTGAGCAAGAGAGAGAGACAGAAATTATTACCGACACACAGAGACTGGGAAAC 495  
 QY 91 ArgTyrSerLysIleAlaLysHisLeuProGlyArgTyrAspAsnGluIleLysAsnTyr 110  
 Db 496 AGATGGCAGAAATGCGAAAGCTACTGCTGGACGAACTGATTAATGCAATCAAGAACAC 555  
 QY 111 TyrArgThrArgIleGlnLysHisIleLysGlnAlaGluAsnPheGlnGlnSerSer 130  
 Db 556 TGGAAATTCAATGCGTCGAGAGGTGCAACAG--GAAGGTATTCTGCAGAGACTTCA 612  
 QY 131 AsnAsnSerGlu-----IleAsnAspHisGlnAlaSerThrSerHisValSerThrMet 148  
 Db 613 AAAGCCAGCCAGCAGAGAGGTGGCCACAGCTTCCAGAAAGACATGATTTAGGCTTT 672  
 QY 149 AlaGluProMetGluMetTyrSerProProCysTyrGlnGlyMetLeuGluProPheSer 168  
 Db 673 GCTCAG-----GCTCCGCTTACAGCTCA-----CTCCCTGGCACC 708  
 QY 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsn 188  
 Db 709 GGCCAG--CCCACTGTTAACACAGCATATTCTATTACCAATTCTGAAGACAAAT 765  
 QY 189 IleAsn 190  
 Db 766 GTCTCC 771

RESULT 8

US-09-167-322-10

Sequence 10. Application US/09167322

Patent No. 6365151

GENERAL INFORMATION:

APPLICANT: Allegheny University of the Health Sciences, Halpern, Michael S.

England, James M.

TITLE OF INVENTION: CANCER VACCINE

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.

STREET: Suite 1800, Two Penn Center Plaza

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19102

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/167,322

FILING DATE: 07-Oct-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US97/00582

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Monaco, Daniel A.

REGISTRATION NUMBER: 30,480

REFERENCE/DOCKET NUMBER: 7933-33 PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-8383

TELEFAX: (215) 568-5549

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1897 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 10:

## Alignment Scores:

Pred. No.: 2.04e-26 Length: 1897  
 Score: 296.00 Matches: 73  
 Percent Similarity: 53.47% Conservative: 35  
 Best Local Similarity: 36.14% Mismatches: 55  
 Query Match: 26.01% Indels: 40  
 DB: 4 Gaps: 9

US-10-021-811-36 (1-206) x US-09-167-322-10 (1-1897)

QY 12 AspProGluValArgLysGlyProThrPheMetGluGluAspLeuIleLeuMetAsnTyr 31  
 Db 233 AACCCAGAACTTAACAAGAGTCAATGACCTAAGAGAGAGATCAAGGTAAATACAAC 292  
 QY 32 TleAlaAsnHisGlyGluGlyValTyrAsnSerLeuAlaLysAlaGlyLeuLys--- 50  
 Db 293 GTCGAGAAATATGCGCAAGCGCTGGTGGACATGCTTAAG-----CAATTGAAGGA 346  
 QY 51 ArgAsnGlyLysSerCysArgLeuArgTyrPheAsnTyrLeuArgProAspValArgArg 70  
 Db 347 AGGATTGGAAAACAGTGCAGAGAGGTGGCACAACCATCTGAATCCAGAGTGAAGAA 406  
 QY 71 GlyAsnIleThrProGluGluGlnLeuLeuIleMetGluLeuHisAlaLysTyrGlyAsn 90  
 Db 407 ACCTCTGAGCAGAGAGAGAGATGAATTATTACAGGACACACAGAGACTGGGAAC 466  
 QY 91 ArgTyrSerLysIleAlaLysHisLeuProGlyArgTyrThrAspAsnGluIleLysAsnTyr 110  
 Db 467 AGATGGCAGAAATGCAAAAGTTCGCTGAGACGAGCTGAATACGCTGTCAAGAACAC 526  
 QY 111 TyrArgThrArgIleGlnLysHisIleLysGlnAlaGluAsnPheGlnGlnSerSer 130  
 Db 527 TGAATTTCACCATCGCGCGGAAGGTGAGCAG--GAGGGTTACCCGACAGAGTCTCC 583  
 QY 131 Asn-----AsnSerGluIleAsnAspHisGlnAlaSerThrSerHisValSerThr 147  
 Db 584 AAAGCCGCGCGCTGCGCAACACCGGCTTCCAGAAAGC--AGCATGTATGAGCC 640  
 QY 148 MetAlaGluProMetGluMetTyrSerProProCys----- 159  
 Db 641 TTGGC-----CACACCCACTGACAGGCCCGCGGGGGCGGCGCAG 685  
 QY 160 -----TyrGlnLysMetLeuGluProPheSerThr----- 169  
 Db 686 GCCCTCTGGGAGTAGTACCTACCTATTACCAATGTGTGAGCCACAAATGCTCCTGCT 745  
 QY 170 GlnPhePro-----ThrIleAsnProAspGlnSerSerCys 181  
 Db 746 CAGATCCCATATCCAGTAGCAGCTGATTAATATTATCAATGTTCTCTCAACC-AGCTGC 804

Qy 182 CysThr 183  
Db 805 TGCAGC 810

RESULT 9  
US-09-402-929-5  
Sequence 5, Application US/09402929  
GENERAL INFORMATION:  
APPLICANT: Temple University - Of The Commonwealth System of Higher Education  
PATENT NO. 6410825  
APPLICANT: Toscani, Antonio  
APPLICANT: Hatton, Kimi  
APPLICANT: Reddy, E. P.  
TITLE OF INVENTION: A-myb NULL MUTANT TRANSGENIC ANIMALS AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.  
STREET: Suite 1800 Two Penn Center Plaza  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19102

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/402,929  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US98/06896  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 6056-214 PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4880 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-402-929-5

Alignment Scores:  
Pred. No.: 26-25 Length: 4880  
Score: 293.00 Matches: 70  
Percent Similarity: 56.22% Conservative: 34  
Best Local Similarity: 37.84% Mismatches: 57  
Query Match: 25.75% Indels: 24  
DB: Gaps: 8

US-10-021-811-36 (1-206) x US-09-402-929-5 (1-4880)

Qy 12 AspProGluValArgLysGlyProTyrThrMetGluGluAspLeuIleLeuMetAsnTyr 31  
Db 348 AATCCTGAATTGATTAAGGGCTCTTGACCTAAGAGAGATGACAGGGTTATTGAATTA 407

Qy 32 IleAlaAsnHisGlyGluGlyValITrPAsnSerLeuAlaLysAlaAlaGlyLeuLys--- 50  
Db 408 GTTACAGAAATATGAGCCCAAAAAGATGCTTTAATTGCAAAA-----CAATTAAAGGA 461

Qy 51 ArgAsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArgArg 70  
Db 462 AGAATAGGCAAGCACTGTAGAGAAAGATGAGCATTAATCTCGAATCTCGAGGTAAAGAA 521

Qy 71 GlyAsnIleThrProGluGluGlnLeuLeuIleMetGluLeuHisAlaLysTrpGlyAsn 90

Db 522 TCTTCCTGAGCAAGAGAGAGAGACAGATCATATGATTAAGACATTAACCGGTGGAAAT 581

Qy 91 ArgTrpSerLysIleAlaLysHisIleuProGlyArgTrpAspAsnGluIleLysAsnTyr 110  
Db 582 CGTTGGCAGAAATATGCCAAACTACTTCGAGAGAGACTGATTAATTCATCAAAATATCAT 641

Qy 111 TrpArgThrArgIleGlnLysHisIleLysGlnAlaGluAsnPheGlnGln-----Gln 128  
Db 642 TGGAAATTCATCATGCAAGAAAGATGAGACAGAGGGCTATTTCAGAGATGAATAA 701

Qy 129 SerSerAsnAsnSerGluIleAsnAspHisGln-----AlaSerThrSerHisValSer 146  
Db 702 TCAGACAGATCTTCATCTTAACCTTCACACAAACCTGTGACGATGATGATATGCA 761

Qy 147 ThrMetAlaGluProMetGluMetTyrSerPro-----ProCysTyrGlnLysMet 163  
Db 762 ACCCAG-----AATCAGTTTACATACCTGTTTCAGATCCCTGGGTATCAGATATG 812

Qy 164 -----LeuGluProPheSerThr-----GlnPhePro 172  
Db 813 TCACCTGAGGCAATGTATAGACATGTTCAGCCTACTTCGCTTTATTCAGCAACCC 872

Qy 173 ThrIleAsnProAsp 177  
Db 873 TTCATTGATGAAGAT 887

RESULT 10  
US-09-402-929-1  
Sequence 1, Application US/09402929  
PATENT NO. 6410825  
GENERAL INFORMATION:  
APPLICANT: Temple University - Of The Commonwealth System of Higher Education  
APPLICANT: Toscani, Antonio  
APPLICANT: Hatton, Kimi  
APPLICANT: Reddy, E. P.  
TITLE OF INVENTION: A-myb NULL MUTANT TRANSGENIC ANIMALS AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.  
STREET: Suite 1800 Two Penn Center Plaza  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19102

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/402,929  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US98/06896  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 6056-214 PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3602 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-402-929-1

## Alignment Scores:

Pred. No.: 1,92e-25 Length: 3602  
 Score: 291.50 Matches: 64  
 Percent Similarity: 59.49% Conservative: 30  
 Best Local Similarity: 40.51% Mismatches: 51  
 Query Match: 25.62% Indels: 13  
 DB: 4 Gaps: 6

US-10-021-811-36 (1-206) x US-09-402-929-1 (1-3602)

OY 12 AspProGluValArgLysGlyProTrpThrMetGluGluAspLeuIleuMetAsnTyr 31  
 DB 500 AATCCGAATTATATAGAGGCTCTTGACTAAGAGAAAGATCAAGAGGTTATTAATTA 559  
 OY 32 IleAlaAsnHisGlyGluGlyValTTPAsnSerLeuAlaLysAlaLysLeuLys--- 50  
 DB 560 GTTCAGAAATATAGGCCAAGAGAGGCTCTTAATTAATTCAGAAA-----CATTTAAAGCA 613  
 OY 51 ArgAsnGlyLysSerCysArgLeuArgTTPLeuAsnTyrLeuArgProAspValArgArg 70  
 DB 614 AATATAGGCAAGCAGTGCAGAGAAAGATGCGCAATCACTGAACTGAACTGAAAGAG 673  
 OY 71 GlyAsnIleThrProGluGluGluLeuIleMetGluLeuHisAlaLysTrpGlyAsn 90  
 DB 674 TCTTCTCGACAGAAAGAGAGAGAGATCATATATGAAGCAGACAGCGCTGGGAAAC 733  
 OY 91 ArgTTPSerLysIleAlaLysHisLeuProGlyArgTTPAspAsnGluIleLysAsnTyr 110  
 DB 734 CTTGGGCGGAGATTGCTTAAGTACTTCTTGAGAGAGATGATTAATTCATCAAAAATCAT 793  
 OY 111 TTPArgThrArgIleGlnLysHisIleLysGlnAlaGluAsnPhgGln-----Gln 128  
 DB 794 TGGAAATTCACCATCGAAGAAAGTGAACAGAGGCTATTTCAGAGATGATGAATAAA 853  
 OY 129 SerSerAsnAsnSerGluIleAsnAspHisGln-----AlaSerThrSerHisValSer 146  
 DB 854 TCAGAGCGGCTCTTCATCAAAACTTCACAGCAAACTTGCGACATATGAGACCATTTGCA 913  
 OY 147 ThrMetAlaGluProMetGluMetCysPro-----ProCysTyrGln 161  
 DB 914 ACCCAG-----AATCAGTTTATCATTTCTGTTCAAGATCCCTGGGTATCAG 958

RESULT 11  
 US-09-402-929-4  
 ; Sequence 4, Application US/09402929  
 ; Patent No. 6410925  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Temple University - Of The Commonwealth System of Higher Education  
 ; APPLICANT: Toscani, Antonio  
 ; APPLICANT: Hatton, Kimi  
 ; APPLICANT: Reddy, E. P.  
 ; TITLE OF INVENTION: A-myb NULL MUTANT TRANSGENIC ANIMALS AND  
 ; TITLE OF INVENTION: USES THEREOF  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Seidel, GONDA, LAVORGNA & MONACO, P. C.  
 ; STREET: Suite 1800 Two Penn Center Plaza  
 ; CITY: Philadelphia  
 ; STATE: PA  
 ; COUNTRY: U.S.A.  
 ; ZIP: 19102  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/402,929  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US98/06896  
 ; FILING DATE:

## ATTORNEY/AGENT INFORMATION:

NAME: Monaco, Daniel A.  
 REGISTRATION NUMBER: 30,480  
 REFERENCE/DOCKET NUMBER: 6056-214 PC  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (215) 568-8383  
 TELEFAX: (215) 568-5549  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 6775 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear

US-09-402-929-4

## Alignment Scores:

Pred. No.: 5.01e-25 Length: 6775  
 Score: 291.50 Matches: 64  
 Percent Similarity: 59.49% Conservative: 30  
 Best Local Similarity: 40.51% Mismatches: 51  
 Query Match: 25.62% Indels: 13  
 DB: 4 Gaps: 6

US-10-021-811-36 (1-206) x US-09-402-929-4 (1-6775)

OY 12 AspProGluValArgLysGlyProTrpThrMetGluGluAspLeuIleuMetAsnTyr 31  
 DB 3749 AATCCGAATTATATAGAGGCTCTTGACTAAGAGAAAGATCAAGAGGTTATTAATTA 3808  
 OY 32 IleAlaAsnHisGlyGluGlyValTTPAsnSerLeuAlaLysAlaLysLeuLys--- 50  
 DB 3809 GTTCAGAAATATAGGCCAAGAGAGGCTCTTAATTAATTCAGAAA-----CATTTAAAGCA 3862  
 OY 51 ArgAsnGlyLysSerCysArgLeuArgTTPLeuAsnTyrLeuArgProAspValArgArg 70  
 DB 3863 AGATAGGCAAGCAGTGCAGAGAAAGATGCGCAATCACTGAACTGAACTGAAAGAG 3922  
 OY 71 GlyAsnIleThrProGluGluGluLeuIleMetGluLeuHisAlaLysTrpGlyAsn 90  
 DB 3923 TCTTCTCGACAGAAAGAGAGAGATCATATATGAAGCAGACAGCGCTGGGAAAC 3982  
 OY 91 ArgTTPSerLysIleAlaLysHisLeuProGlyArgTTPAspAsnGluIleLysAsnTyr 110  
 DB 3983 CTTGGGCGGAGATTGCTTAAGTACTTCTTGAGAGAGATGATTAATTCATCAAAAATCAT 4042  
 OY 111 TTPArgThrArgIleGlnLysHisIleLysGlnAlaGluAsnPhgGln-----Gln 128  
 DB 4043 TGGAAATTCACCATCGAAGAAAGTGAACAGAGGCTATTTCAGAGATGATGAATAAA 4102  
 OY 129 SerSerAsnAsnSerGluIleAsnAspHisGln-----AlaSerThrSerHisValSer 146  
 DB 4103 TCAGAGCGGCTCTTCATCAAAACTTCACAGCAAACTTGCGACATATGAGACCATTTGCA 4162  
 OY 147 ThrMetAlaGluProMetGluMetCysPro-----ProCysTyrGln 161  
 DB 4163 ACCCAG-----AATCAGTTTATCATTTCTGTTCAAGATCCCTGGGTATCAG 4207

RESULT 12  
 US-08-306-691B-46  
 ; Sequence 46, Application US/08306691B  
 ; Patent No. 5734039  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Calabretta, Bruno  
 ; APPLICANT: Skorski, Tomasz  
 ; TITLE OF INVENTION: ANTISENSE  
 ; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES  
 ; NUMBER OF SEQUENCES: 55  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P. C.  
 ; STREET: Two Penn Center, Suite 1800  
 ; CITY: Philadelphia  
 ; STATE: Pennsylvania  
 ; COUNTRY: U.S.A.

ZIP: 19102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/306,691B  
FILING DATE: September 15, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 8321-8  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
TELEX: No. 5734039e  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2638 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-306-691B-46

Alignment Scores:  
Pred. No.: 3.37e-23 Length: 2638  
Score: 271.50 Matches: 67  
Percent Similarity: 52.17% Conservative: 29  
Best Local Similarity: 36.41% Mismatches: 59  
Query Match: 23.86% Indels: 29  
DB: Gaps: 8

US-10-021-811-36 (1-206) x US-08-306-691B-46 (1-2638)

Qy 5 GlnGlnCysLys-----ThSerGlnAAspProGluValArgLysGlyProTir 20  
Db 326 CAGCATCCACAGAGTGGCTGAGAGCTTTTGATCCAGCTTGTCAGAGGCGCATGG 385  
Qy 21 ThrMetGluGluAspLeuIleuMetAsnTyrIleAlaAsnHisGlyGluGlyValTir 40  
Db 386 ACCAAGAGGAAGACCAAAAGTCACTGAGCTGTAAAGAGTATGGACAAACAGCTGG 445  
Qy 41 AsnSerIeuAlaLysAlaAlaGlyLeuLys--ArgAsnGlyLysSerCysArgLeuArg 59  
Db 446 ACACTGATTGCCAAG-----CACCTGAAGGGCGGCTGGGGAAGCAGTGCCTGAACGC 499  
Qy 60 TrpLeuAsnTyrLeuArgProAspValAlaArgArgGlyAsnIleThrProGluGluGlnLeu 79  
Db 500 TGGCAACAACCTCAACCTGAGGTGAAGAGTCTTGCTGACCGAGAGAGAGACCGC 559  
Qy 80 LeuIleMetGluLeuHisAlaLysTirpGlyAsnArgTirpSerLysIleAlaLysHisIleu 99  
Db 560 ATCATCTCGAGGCCCAACAAGTGTGCGCAACCCCTGCGCGAGATGCCAAGATGTG 619  
Qy 100 ProGlyArgThrAspAsnGluIleLysAsnTyrTirpArgThrArgGlyGlnLysHisIle 119  
Db 620 CCAGGAGAGACAGACAACTGCTGTGAAGATCACTGAACCTACATCAAAAGAGAGTGT 679  
Qy 120 LysGlnAlaGluAsnPhnGlnGlnGlnSerAsnAsnSer----- 133  
Db 680 ---GACACAGAGAGCTTCTTGAGCGAGTCCAAAGCTCCAAAGCCCAAGTACTTGTCTG 736  
Qy 134 ---GluIleAsnAspHisGlnAlaSerThrSerHisValSerThrMetAlaGluProMet 152  
Db 737 CTGAGCTCGAGGACAAAGAGCGCTCCAGAGT-----GCCAGCCCAAG 781  
Qy 153 GluMetTyrSerProProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPhePro 172

Db 782 GAA-----GCCACGAGGAAGTCTT-----CTGACCACTGGGCC 814  
Qy 173 ThrIleAsnPro 176  
Db 815 TCCGTCCCTCT 826

RESULT 13  
US-08-485-139-1  
Sequence 1, Application US/08485139  
Patent No. 5880331  
GENERAL INFORMATION:  
APPLICANT: KREBBERS, Emno  
APPLICANT: WILLIAMS, Mark  
TITLE OF INVENTION: USE OF ANTHOCYANIN GENES TO MAINTAIN  
TITLE OF INVENTION: MALE STERILE PLANTS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,139  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 010830-096  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4059 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: C1 gene of Zea mays  
FEATURE:  
NAME/KEY: -  
LOCATION: 279..284  
OTHER INFORMATION: /label= HpaI  
FEATURE:  
NAME/KEY: -  
LOCATION: 447..452  
OTHER INFORMATION: /label= EcoRI  
FEATURE:  
NAME/KEY: -  
LOCATION: 1735..1740  
OTHER INFORMATION: /label= AacII  
FEATURE:  
NAME/KEY: -  
LOCATION: 1505..1510  
OTHER INFORMATION: /label= EcoRI  
FEATURE:  
NAME/KEY: -  
LOCATION: 2081..2086  
OTHER INFORMATION: /label= XhoI  
FEATURE:  
NAME/KEY: -  
LOCATION: 2418..2430  
OTHER INFORMATION: /label= SfiI









GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 8, 2003, 11:08:55 ; Search time 60 Seconds  
(without alignments)  
1627.410 Million cell updates/sec

Title: US-10-021-811-36

Perfect score: 1138

Sequence: 1 MDKQOQCKTSODEVRKGPW.....NNINYSMEDSWMSQLLNGD 206

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPRO.spool/US10021811/runat\_03022003\_111222\_25671/app\_query.fasta\_1.391  
-DB=Published Applications NA -OPM=fasta -SUFFIX=rmpb -MINMATCH=0.1  
-LOOPEL=0 -IOGEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62  
-TRANS=human40.cdd -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10021811@cgn\_1.1.33 @runat\_03022003\_111222\_25671  
-NCPU=6 -ICPU=3 -NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1138	100.0	863	9	US-10-021-811-35
2	1062	93.3	988	9	US-10-021-811-29
3	874.5	76.8	805	9	US-10-021-811-37
4	829.5	72.9	910	9	US-10-021-811-33

5	609	53.5	382	10	US-09-878-574-1305	Sequence 1305, Ap
6	591	51.9	330	10	US-10-021-811-31	Sequence 31, Appl
7	587.5	51.6	392	10	US-09-878-574-33	Sequence 33, Appl
8	586	51.5	438	10	US-09-878-574-5021	Sequence 5021, Ap
9	562	49.4	375	10	US-09-878-574-327	Sequence 327, App
10	560.5	49.3	379	10	US-09-878-574-3147	Sequence 3147, Ap
11	556.5	48.9	350	10	US-09-878-574-50	Sequence 50, Appl
12	545.5	47.9	365	10	US-09-878-574-194	Sequence 194, App
13	528	46.4	1236	9	US-10-021-811-45	Sequence 45, Appl
14	519.5	45.7	587	9	US-10-021-811-17	Sequence 17, Appl
15	518.5	45.6	1181	9	US-10-021-811-47	Sequence 47, Appl
16	508.5	44.7	1317	9	US-10-021-811-9	Sequence 9, Appl1
17	505	44.4	500	9	US-10-021-811-41	Sequence 41, Appl1
18	503.5	44.2	1348	9	US-10-021-811-43	Sequence 43, Appl
19	502.5	44.2	557	9	US-10-021-811-27	Sequence 27, Appl
20	494.5	43.5	352	10	US-09-878-574-2390	Sequence 2390, Ap
21	494	43.4	771	9	US-10-021-811-1	Sequence 1, Appl1
22	490	43.1	1186	9	US-10-021-811-49	Sequence 49, Appl
23	485	42.6	751	9	US-10-021-811-39	Sequence 39, Appl
24	475	41.7	822	9	US-09-938-842A-2373	Sequence 2373, Ap
25	464	40.8	786	9	US-09-938-842A-2294	Sequence 2294, Ap
26	459	40.3	810	9	US-09-938-842A-1293	Sequence 1293, Ap
27	447	39.3	640	9	US-10-021-811-21	Sequence 21, Appl
28	438	38.5	521	9	US-10-021-811-59	Sequence 59, Appl
29	419	36.8	708	9	US-09-938-842A-1292	Sequence 1292, Ap
30	418.5	36.8	293	10	US-09-878-574-2675	Sequence 2675, Ap
31	415.5	36.5	771	9	US-09-938-842A-2242	Sequence 2242, Ap
32	410	36.0	960	9	US-09-938-842A-2138	Sequence 2138, Ap
33	400.5	35.2	774	9	US-09-938-842A-2636	Sequence 2636, Ap
34	394	34.6	782	9	US-10-021-811-3	Sequence 3, Appl1
35	393.5	34.6	1123	9	US-10-021-811-13	Sequence 13, Appl
36	391	34.4	514	9	US-10-008-118A-5	Sequence 5, Appl1
37	391	34.4	514	10	US-09-443-704-5	Sequence 5, Appl1
38	390.5	34.3	741	9	US-09-938-842A-121	Sequence 121, App
39	387.5	34.1	601	9	US-10-021-811-5	Sequence 5, Appl1
40	387	34.0	1141	9	US-10-008-118A-11	Sequence 11, Appl
41	387	34.0	1141	10	US-09-443-704-11	Sequence 11, Appl
42	385	33.8	849	9	US-09-938-842A-2082	Sequence 2082, Ap
43	384	33.7	1168	9	US-10-008-118A-23	Sequence 23, Appl
44	384	33.7	1168	10	US-09-443-704-23	Sequence 23, Appl
45	382.5	33.6	1044	10	US-09-770-445-167	Sequence 167, App

## ALIGNMENTS

RESULT 1  
US-10-021-811-35  
; Sequence 35, Application US/10021811  
; Publication No. US20030024007A1  
; GENERAL INFORMATION:  
; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Fang, Yiwen  
; APPLICANT: Odell, Joan  
; APPLICANT: Weng, Zude  
; TITLE OF INVENTION: Plant Myb Transcription Factor Homologs  
; FILE REFERENCE: B01294 US NA  
; CURRENT APPLICATION NUMBER: US/10/021,811  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: 60/110,609  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 35  
; LENGTH: 863  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-10-021-811-35

Alignment Scores:  
Pred. No.: 4.9e-127 Length: 863  
Score: 1138.00 Matches: 206  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-021-811-36 (1-206) x US-10-021-811-35 (1-863)

```

QY 1 MetApLySLVSGInGInCysLysThrSerGlnAspProGluValArgLysGlyProTyr 20
DB 29 ATGGATTAATAAACAACAGCTGTAGACGCTCTCAAGATCCTGAAAGTGAAGAAAGGAGCTTTGG 88
QY 21 ThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHISGlyGluGlyValTyr 40
DB 89 ACAATGGAAGAAGACTGATCTTGATGAACTATTTGCAATTCATGGGGAAGGTGTTGG 148
QY 41 AunSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTyr 60
DB 149 AACTCTTTGGCCAAAGCTGTGCTGTCTCAACGTAACGGAAGAGATTGCCGCTAAGGTGG 208
QY 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
DB 209 CTAAATTAACCTCCGCTCGATGTTAGAAAGGGAATATTACACCCGGAACAACCTTTTG 268
QY 81 IleMetGluLeuHISAlaLysTyrGlyAsnArgTyrSerLysIleAlaLysHISLeuPro 100
DB 269 ACTATGGAGCTTCACGCAAGGGGGAACAGGTGTCTCAAAATTTGCCAAGCATCTACCT 328
QY 101 GlyArgThrAspAsnGluIleLysAsnTyrTyrArgThrArgIleGlnLysHISLys 120
DB 329 GAAAGACTGATTAATGAGATCAAGAACTATTGAGAGCAAGATCCAGAAACATCATCAG 388
QY 121 GlyArgThrAspAsnGluIleLysAsnTyrTyrArgThrArgIleGlnLysHISLys 140
DB 389 CAAAGTGAACTTTCACCAACAGAGTAGTATATTTGTGATPAATAGATCACCAAGCT 448
QY 141 SerThrSerHISValSerThrMetAlaGluProMetGluMetTyrSerProCysTyr 160
DB 449 AGCACTACCACTGTTTCCACCATGGCTGAGCCCATGAGATGATTTCTCCACCTGTTAT 508
QY 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSer 180
DB 509 CAAAGAAATGTAGAGCCATTTCACATCACTGATCCCTCAATTAATCCGATCAATCCAGT 568
QY 181 CysCysThrAsnAspAsnAsnAsnIleAsnTyrTyrSerMetGluAspSerTyrSerMet 200
DB 569 TTTGTATACCAATGACAAACAAACATTAATGAGACATGAGATGAGTATGCTGTCATAG 628
QY 201 GlnLeuLeuAsnGlyAsp 206
DB 629 CAATTACTGAACGGTGAT 646

```

#### RESULT 2

```

US-10-021-811-29
; Sequence 29, Application US/10021811
; Publication No. US20030024007A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Fang, Yiwen
; APPLICANT: Ode11, Joan
; APPLICANT: Meng, Zude
; TITLE OF INVENTION: Plant Myb Transcription Factor Homologs
; FILE REFERENCE: BB1294 US NA
; CURRENT APPLICATION NUMBER: US/10/021.811
; PRIORITY FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: 60/110,609
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 988
; TYPE: DNA
; ORGANISM: Glycine max
US-10-021-811-29
Alignment Scores:

```

Pred. No.: 7,44e-118 Length: 988  
Score: 1062.00 Matches: 197  
Percent Similarity: 95.67% Conservative: 2  
Best Local Similarity: 94.71% Mismatches: 7  
Query Match: 93.32% Indels: 2  
DB: 9 Gaps: 2

US-10-021-811-36 (1-206) x US-10-021-811-29 (1-988)

```

QY 1 MetApLySLVSGInGInCysLysThrSerGlnAspProGluValArgLysGlyProTyr 20
DB 61 ATGGATTAATAAACAACAGCTGTAGACGCTCTCAAGATCCTGAAAGTGAAGAAAGGAGCTTTGG 120
QY 21 ThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHISGlyGluGlyValTyr 40
DB 121 ACGATGGAAGAAGACTGATCTTGATGAACTATTTGCAATTCATGGGGAAGGTGTTGG 180
QY 41 AunSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTyr 60
DB 181 AATCTTTGGCCAAAGCTGTGCTGTCTCAACGTAACCGGAAGAGATTGCCGCTAAGGTGG 240
QY 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
DB 241 CTAAATTAACCTCCGCTCGATGTTAGAAAGGGAATATTACACCCGGAACAACCTTTTG 300
QY 81 IleMetGluLeuHISAlaLysTyrGlyAsnArgTyrSerLysIleAlaLysHISLeuPro 100
DB 301 ATCATGAGACTTCACGCAAGGGGGAACAGGTGTCTCAAAATTTGCCAAGCATCTACCT 360
QY 101 GlyArgThrAspAsnGluIleLysAsnTyrTyrArgThrArgIleGlnLysHISLys 120
DB 361 GGTAGACAGATTAATGAGATCAAGAACTATTGAGAGCAAGATCCAGAACATCATCAG 420
QY 121 GlnAlaGluAsnPheGlnGlnInSerSerAsnAsnSerGluIleAsnAspHISGlnAla 140
DB 421 CAAGTGAAACTTTCAGCAACAAATCAATCACTGATPAATAGATCACCAACT 480
QY 141 SerThrSerHISValSerThrMetAlaGluProMetGluMetTyrSerProCysTyr 160
DB 481 AGCACTACCACTGTTTCCACCATGGCTGAGCCCATGAGATGATTTCTCCACCTGTTAT 540
QY 161 GlnGlyMetLeuGluProPheSerThr---GlnPheProThrIleAsnProAspGlnSer 179
DB 541 CAAAGAAATGTAGAGCCATTTCACATCACTGATCCCTCAATTAATCCGATCAATCCAGT 600
QY 180 SerCysCysThrAsnAspAsnAsn---IleAsnTyrTyrSerMetGluAspSerTyr 198
DB 601 AGTTGTTGACCAATGACAAACAAACATTAATGAGACATGAGATGAGTATGCTGG 660
QY 199 SerMetGlnLeuLeuAsnGlyAsp 206
DB 661 TCATATGCACTTACTGAACGGGGAT 684

```

#### RESULT 3

```

US-10-021-811-37
; Sequence 37, Application US/10021811
; Publication No. US20030024007A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Fang, Yiwen
; APPLICANT: Ode11, Joan
; APPLICANT: Meng, Zude
; TITLE OF INVENTION: Plant Myb Transcription Factor Homologs
; FILE REFERENCE: BB1294 US NA
; CURRENT APPLICATION NUMBER: US/10/021.811
; PRIORITY FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: 60/110,609
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 37
; LENGTH: 805
; TYPE: DNA

```

```
; ORGANISM: Glycine max
US-10-021-811-37

Alignment Scores:
Pred. No.: 1,51e-95      Length: 805
Score: 874.50           Matches: 168
Percent Similarity: 85.51%  Conservative: 9
Best Local Similarity: 81.16%  Mismatches: 23
Query Match: 76.85%      Indels: 7
DB: 9                   Gaps: 5

US-10-021-811-36 (1-206) x US-10-021-811-37 (1-805)

Qy 4 LysGlnGlnCys---LysThrSerGlnAspProGluValArgLysGlyProTrrPThrMet 22
    |||::|
Db 1 AAAAAACCATGCACTCATCTCATCTCATGATCTCTCAAGTGAGAAAGGACCAATGACCATG 60

Qy 23 GluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrrPAsnSer 42
    |||::|
Db 61 GAGAGAGACTTGATCTGATTAACCTATATTGCAATCAACGATGAGGATGTTGGAACTCC 120

Qy 43 LeuAlaIysAlaIaGlyLeuLysArgAsnGlyLysSerCysArgLeuAArgTrrPLeuAsn 62
    |||::|
Db 121 TTAGCCAAAGCTTCTGCTGCTCAACGAAAGGAGATGTCGATCTCGCTTAAC 180

Qy 63 TyrLeuAArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuIleMet 82
    |||::|
Db 181 TACCTTCCTCTGATGTTTGAAGAAGAAACATTACCCCGAGAAACAGCTTTTATCATATA 240

Qy 83 GluLeuHisAlaIaLysTrrPGLYAsnAArgTrrPserLysIleAlaLysHisLsLeuProGlyArg 102
    |||::|
Db 241 GAACCTTCATGCAAGAGTGGGCAATAGGTGTCGCAAAATTTGCAAAAGCATCTTCCAGAGAGA 300

Qy 103 ThrAspAsnGluIleLysAsnTyrTrrPArgThrArgIleGlnLysHisIleLysGlnAla 122
    |||::|
Db 301 ACTGCAATGAGATTAAAGACTTCTGAGAAACAAAGATCCAAAGACCAATTAAGCAAGCT 360

Qy 123 GluAsnPhenGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla----- 140
    |||::|
Db 361 GAGACTTCACAACAACAT-----GCTAATTCAAGAAATATATGATCATCAAGCAAGCACT 414

Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProProCysTyr 160
    |||::|
Db 415 AGTACTACCAAGATGTCACCATGTCACATCCAAATGAGACTTCTCTCCACCTCATAC 474

Qy 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSerSer 180
    |||::|
Db 475 CAAGGAACTTTGAGCCATTCACACTCAATTCCTCTACAACTACT--GATCAATCAAGT 531

Qy 181 CysCysThrAsnAspAsnAsnAsnIleAsnTyrTrrPserMetGluAspSerTrrP---Ser 199
    |||::|
Db 532 TGTTTTACCAACCAACGACGACAAACAACATTTGAGACATCGAGATATATGTCGTCT 591

Qy 200 MetGlnLeuLeuAsnGlyAsp 206
    |||::|
Db 592 ATGCAATTACTCAATGAGAT 612

RESULT 4
US-10-021-811-33
; Sequence 33, Application US/10021811
; Publication No. US20030024007A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Fang, Yiwen
; APPLICANT: Odell, Joan
; APPLICANT: Meng, Zude
; TITLE OF INVENTION: Plant Myb Transcription Factor Homologs
; FILE REFERENCE: BB1294 US NA
; CURRENT APPLICATION NUMBER: US/10/021,811
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: 60/110,609
; PRIOR FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 63
```

```
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 33
; LENGTH: 910
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (798)
; NAME/KEY: unsure
; LOCATION: (807)
; NAME/KEY: unsure
; LOCATION: (814)
US-10-021-811-33

Alignment Scores:
Pred. No.: 4.38e-90      Length: 910
Score: 829.50           Matches: 168
Percent Similarity: 83.89%  Conservative: 9
Best Local Similarity: 79.62%  Mismatches: 23
Query Match: 72.89%      Indels: 12
DB: 9                   Gaps: 7

US-10-021-811-36 (1-206) x US-10-021-811-33 (1-910)

Qy 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrrP 20
    |||::|
Db 37 ATGACAAAAA---CCATGCACATCATCTCATGATCTCAAGATGAGAAAGGACCATG 93

Qy 21 ThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrrP 40
    |||::|
Db 94 ATCATGAAAGAAAGACTTGATCTTGATTAACCTATATTGCAAAATCAACGATGAGGATGTTGG 153

Qy 41 AsnSerLeuAlaLysAlaIaGlyLeuLysArgAsnGlyLysSerCysArgLeuAArgTrrP 60
    |||::|
Db 154 AATCTTTAGCCAAAGCTTCTGCTTAAACGAAAGGAGAAAGATGTCGATCCGTTGG 213

Qy 61 LeuAsnTyrLeuAArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeu 80
    |||::|
Db 214 CTAACTACCTTCGCTGATGTTAGAAAGAGAAACATTACCCCGAGAAACAGCTTTTG 273

Qy 81 IleMetGluLeuHisAlaLysTrrPGLYAsnAArgTrrPserLysIleAlaLysHisLsLeuPro 100
    |||::|
Db 274 ATCATGAACTTCATGCAAGTGGGCAATAGGTGTCGCAAAATTTGCAAAAGCATCTTCCA 333

Qy 101 GlyAArgThrAspAsnGluIleLysAsnTyrTrrPArgThrArgIleGlnLysHisIleLys 120
    |||::|
Db 334 GGAAGAATGACATGATGATTAAGAACTTCTGAGAACTAGATCCAGAAACATTAAG 393

Qy 121 GlnAlaGluAsnPhenGlnGln-----GlnSerSerAsnAsnSerGluIleAsnAspHis 138
    |||::|
Db 394 CAAGCTGAGACTTCACACAACAACATGTAATTCATCAGAGAAATAGTAAT--AATGATCAT 450

Qy 139 GlnAla-----SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSer 156
    |||::|
Db 451 CAAGCAAGCAATGACACTGACCAAGGTGTCACATGCGACATCCAAATATAGACTTTCCT 510

Qy 157 ProProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnPro 176
    |||::|
Db 511 TCACCTCATACCAACAACACTTTTGAAGCATTTCA-ACCTCAATTCCTCAATCAAT--- 566

Qy 177 AspGlnSerSerCysCysThrAsnAspAsnAsnAsnIleAsnTyrTrrPserMetGluAsp 196
    |||::|
Db 567 GATCAATCAAGATTGTTGTCACGCAACAAC-----AACTATTGAGATTCGAGAT 617

Qy 197 SerTrrP---SerMetGlnLeuLeuAsnGlyAsp 206
    |||::|
Db 618 ATCTGTCGTCTATGCAATTACTCAATGAGAT 650

RESULT 5
US-09-878-574-1305
; Sequence 1305, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO: 1305
; LENGTH: 382
; TYPE: RNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB028-039-Q1-B1-G9
US-09-878-574-1305

Alignment Scores:
Pred. No.: 2,96e-64 Length: 382
Score: 609.00 Matches: 110
Percent Similarity: 93.60% Conservative: 7
Best Local Similarity: 88.00% Mismatches: 8
Query Match: 53.51% Indels: 0
DB: 10 Gaps: 0

US-10-021-811-36 (1-206) x US-09-878-574-1305 (1-382)

QY 10 SerGlnAEPProGluValArgIysGlyProThrPthMetGluGluValAspLeuMet 29
Db 2 TCTCATGATCCGAAAGTGAAGGAGGACCATGATCATGGAAGAACTTGATCTGATA 61
QY 30 AsnTyrIleAlaAsnHISGlyGluGlyValTyrPAsnSerIleuAlaIleGlyLeu 49
Db 62 AACATATATGGAAATTCACGGTGAAGGTGTGGAAATCTTACGCAAAAGCTTGCTT 121
QY 50 LysArgAsnGlyLysSerCysArgLeuArgTyrLeuAsnTyrLeuAspProAspValArg 69
Db 122 AAAGGAACGGGAAAGAGTGTGCACTCCGTTGGCTTAACTTCTTCCTGATGTTAGA 181
QY 70 ArgGlyAsnIleThrProGluGluGlnLeuLeuIleMetGluLeuHISAlaIysTyrGly 89
Db 182 AAGAGAAACATTACACCCGGAAGAACACCTTTGATCATAGAACTTCATGCAAAAGTGGGC 241
QY 90 AsnArgTyrSerIleAlaIleAlaIleAlaIleLeuProGlyArgTyrAspAsnGluIleYAsn 109
Db 242 AATAGGTGCTCAAAATTCGAAAGCATCTTCCAGGAAGACTGCAATGAGATTAGAAC 301
QY 110 TyrThrArgThrArgGlyGlnIleYAsnHISIleYsGlnAlaGluAsnIleGlnGlnSer 129
Db 302 TTCTGGAAACTAGGATTCGAAAGCATTAAGCAAGCTGAGACTTCACAAACAATGGT 361
QY 130 SerAsnAsnSerGlu 134
Db 362 AATTCATCCAGAA 376

RESULT 6
US-10-021-811-31
; Sequence 31, Application US/10021811
; Publication No. US20030024007A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Fang, Yiwon
; APPLICANT: Odell, Joan
; APPLICANT: Weng, Zude
; TITLE OF INVENTION: Plant Myb Transcription Factor Homologs
; FILE REFERENCE: BB1294 US NA
; CURRENT APPLICATION NUMBER: US/10/021,811
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: 60/110,609
; PRIOR FILING DATE: 1998-December-02
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Microsoft Office 97
```

```
; SEQ ID NO 31
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (301)
; NAME/KEY: unsure
; LOCATION: (356)
; NAME/KEY: unsure
; LOCATION: (388)
; NAME/KEY: unsure
; LOCATION: (393)
; NAME/KEY: unsure
; LOCATION: (470)
; NAME/KEY: unsure
; LOCATION: (485)
; NAME/KEY: unsure
; LOCATION: (495)
; NAME/KEY: unsure
; LOCATION: (513)
; NAME/KEY: unsure
; LOCATION: (517)
; NAME/KEY: unsure
; LOCATION: (522)
US-10-021-811-31

Alignment Scores:
Pred. No.: 6.84e-62 Length: 530
Score: 591.00 Matches: 122
Percent Similarity: 83.33% Conservative: 8
Best Local Similarity: 78.21% Mismatches: 25
Query Match: 51.93% Indels: 3
DB: 9 Gaps: 1

US-10-021-811-36 (1-206) x US-10-021-811-31 (1-530)

QY 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProThr 20
Db 7 ATGGACAAGAAAGCTTGGC---AACACGTCTCATGATCTCGAAGGAGAGGGGCGCATGG 63
QY 21 ThrMetGluGluAspLeuIleuMetAsnTyrIleAlaAsnHISGlyGluGlyValTyr 40
Db 64 ACAATGGAAGAAAGCTTATATCTGATCATCTATATTCGCATTCACGGGAAGGGGTTGG 123
QY 41 AsnSerIleuAlaIleAlaIleGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTyr 60
Db 124 AACTCTTTGGCCAAAGGCTGCTGCACTTAAACGTACCGGAAGAGTTCCCGCTCCGGTGG 183
QY 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGlnLeuLeu 80
Db 184 CTAAACTACTCCCGTCCGATGTTAGAAAGGAGATTTACACCCGAGGAACAGCTTTG 243
QY 81 IleMetGluLeuHISAlaIleValTyrGlyAsnArgTyrSerIleAlaIleYsLeuPro 100
Db 244 ATCATGGAACCTTCATGCAAAAGTGGGAAACAGGTGTCCAAATTCGCAAGCATCTTACC 303
QY 101 GlyArgThrAspAsnGluIleYAsnTyrTyrArgThrArgIleGlnYsHISIleYs 120
Db 304 GGAAGGACTGATTAATGAGATTAGAACTACTGAGGCAAGAT-CAAGAAACANCTCAG 362
QY 121 GlnIleGluAsnIleGlnGlnSerSerAsnAsnSerGluIleAsnAspHISGlnAla 140
Db 363 CAAAGCTTCACAA-CTTCACAAACANAGTANTATTTGAGATTAATTCATCCCAAGC 421
QY 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetYrSer 156
Db 422 TTGCACAACCAATTTGCAACATGCGCAACCCAAAAAACATAATCTCA 469

RESULT 7
US-09-878-574-33
; Sequence 33, Application US/09878574
; Patent No. US20020110548A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 33
; LENGTH: 392
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-006-Q1-B1-A8
; US-09-878-574-33

Alignment Scores:
Pred. No.: 1,156-61 Length: 392
Score: 587.50 Matches: 109
Percent Similarity: 91.80% Conservative: 3
Best Local Similarity: 89.34% Mismatches: 9
Query Match: 51.63% Indels: 1
DB: 10 Gaps: 1

US-10-021-811-36 (1-206) x US-09-878-574-33 (1-392)
Qy 1 MetAspLySgInGInCysLySThrSerGInAspProGluValArgLySgLyProTr 20
Db 29 ATGACGAAAGAGCTTGGC--AACACGCTCATATCTCTGAAGTGAAGAAAGGCGCATGG 85
Qy 21 ThrMetGluGluAspLeuIleuMetAsnTyrlleAlaAsnHisgLyGluGlyValTr 40
Db 86 ACAATGGAAGAAAGCTTATCTTGATCACTATATTGGCAATTCACGGGGAAGGGGTTTGG 145
Qy 41 AsnSerLeuAlaLySAlaAlaGlyLeuLySArgAsnGlyLySserCysArgLeuArgTr 60
Db 146 AACTTTTGGCCAGAGCTGCTGACTTAACGATACCGGAAGAGTGGCGGCTCGGTTGG 205
Qy 61 LeuAsnTyrlleuArgProAspValArgArgGlyAsnIleThrProGluGluGlnleu 80
Db 206 CTAACACTACCTCCGCTCGATGTTTGAAGAGGAAATATTACACCGAAGAAAGCTTTTGG 265
Qy 81 IleMetGluLeuHisAlaLySArgTrGlyAsnArgTrpSerLySIIeAlaLySHisleuPro 100
Db 266 ATCATGGAACCTTCATGACAGAGTGGGAAACAGGTGTCCTCAAAATTGCCAAGCATCTACCC 325
Qy 101 GlyArgThrAspAsnGluIleLySAsnTyrlTrpArgThrArgIIeGlnLySHisleu 120
Db 326 GGAAGGACTGATCTGAGATTAAAGAACTACTGAGAGCAATGATCATGATGACCTCAAG 385
Qy 121 GlnAla 122
Db 386 CAAGCT 391

RESULT 8
US-09-878-574-5021
; Sequence 5021, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
```

```

; SEQ ID NO 5021
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-010-Q1-B1-B1
; US-09-878-574-5021

Alignment Scores:
Pred. No.: 2,056-61 Length: 438
Score: 586.00 Matches: 112
Percent Similarity: 87.05% Conservative: 9
Best Local Similarity: 80.58% Mismatches: 15
Query Match: 51.49% Indels: 3
DB: 10 Gaps: 1

US-10-021-811-36 (1-206) x US-09-878-574-5021 (1-438)
Qy 1 MetAspLySgInGInCysLySThrSerGInAspProGluValArgLySgLyProTr 20
Db 29 ATGACGAAAGAGCTTGGC--AACACGCTCATATCTCTGAAGTGAAGAAAGGCGCATGG 88
Qy 20 ThrMetGluGluAspLeuIleuMetAsnTyrlleAlaAsnHisgLyGluGlyValTr 40
Db 89 GACAGGGAAGAAAGACTTGATCTTGATTAACATATGCAAAATCACGGTGAACGTGTTTG 148
Qy 40 PAsnSerLeuAlaLySAlaAlaGlyLeuLySArgAsnGlyLySserCysArgLeuArgTr 60
Db 149 GAACCTCTTAGCCAAAGCTTCTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 208
Qy 60 PleuAsnTyrlleuArgProAspValArgArgGlyAsnIleThrProGluGluGlnleu 80
Db 209 GCTAACTACCTTCGCTCGATGTTTGAAGAGGAAACATTTACACCGGGAAGAGCTTTT 268
Qy 80 uIleMetGluLeuHisAlaLySArgTrGlyAsnArgTrpSerLySIIeAlaLySHisleuPr 100
Db 269 GATCATTAAGAACTTCATGACAAAGTGGGCAATACGTGTCCTCAAAATTGCCAAGCATCTTC 328
Qy 100 ogLyArgThrAspAsnGluIleLySAsnTyrlTrpArgThrArgIIeGlnLySHisleu 120
Db 329 AGGAAGAACTGACAAATGATTAAGAACTTCTGAGAAACAAAGATTCAAAACACATTAA 388
Qy 120 gGlnAlaGluAsnPhelGlnGlnInserSerAsnAsnSerGluIleAsnAspHis 138
Db 389 GCAAGCTGAGACTTCACAAACAT-----GGTAATTCAGAGAAATATGATCAT 437

RESULT 9
US-09-878-574-327
; Sequence 327, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 327
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-052-Q1-B2-D5
; US-09-878-574-327

Alignment Scores:
Pred. No.: 1,216-58 Length: 375
Score: 562.00 Matches: 104
Percent Similarity: 93.28% Conservative: 7
Best Local Similarity: 87.39% Mismatches: 6
```

Query Match: 49.38% Indels: 2  
DB: 10 Gaps: 1  
US-10-021-811-36 (1-206) x US-09-878-574-327 (1-375)  
QY 23 GUGUAAPLeuileuemeCAsnTYrilleaIaasnHISglUGluGValTTPAsnSer 42  
DB 23 GAAGAAAGCTGATCTTGAATTAATATGCAATCAAGGTGGATTGGAATCTCC 82  
QY 43 leuAlaYsaIaAlaGlyLeuLYsaArgasnGlyYsserCYsaGleuAgtTTPLeuasn 62  
DB 83 TTAGCCAAAGCTTGCTGCTCAAGAAAGGGAAGAGTTTCATCTCCGTGGCTAAC 142  
QY 63 TYrleuArGProaspValaArgArgGlyYasnIleThrProglUGluGlnLeuileuMeC 82  
DB 143 TACCTTCCTCGATGTAGAAAGAAACATTAACCCGGAAGAGCTTTATATATA 202  
QY 83 GUGUeuhSAlaYsTTPGlyYasnArgTTPserYsilleaIaYsHISleuProGlyArg 102  
DB 203 GAACCTCAAGCAAGTGGGCAATGAGTGTCCAAATTCGAAAGCATCTTCGAGGAAGA 262  
QY 103 ThrAspAengluileLYsaSnTYrTPArgThArgIleGlnYsHISleYsGlnAla 122  
DB 263 ACTGACAAATGAGATTAAAGAACTTCTGAGAACAGATCCAAAGCATTAAGCAAGCT 322  
QY 123 GUAasnPhengIngnInserSerAsnAsnSerGlnIleasnAepHISGlnAlaSer 141  
DB 323 CAGACTTCACAAAGACAT-----GTTAATTCAGAAATATGATCATCAAGCAAGC 373  
RESULT 10  
US-09-878-574-3147  
/ Sequence 3147, Application US/09878574  
/ Patent Nc. US20020110548A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Byrum, Joseph R.  
/ APPLICANT: La Rosa, Thomas J.  
/ APPLICANT: Thompson, Michael D.  
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
/ TITLE OF INVENTION: Plants  
/ FILE REFERENCE: 38-21 (15401) B  
/ CURRENT APPLICATION NUMBER: US/09/878,574  
/ CURRENT FILING DATE: 2001-12-21  
/ PRIOR APPLICATION NUMBER: 09/333,535  
/ PRIOR FILING DATE: 1999-06-14  
/ NUMBER OF SEQ ID NOS: 15775  
/ SEQ ID NO 3147  
/ LENGTH: 379  
/ TYPE: DNA  
/ ORGANISM: Glycine max  
/ OTHER INFORMATION: Clone ID: LIB3028-015-Q1-B1-E7  
US-09-878-574-3147  
Alignment Scores:  
Pred. No.: 1,86e-58 Length: 379  
Score: 560.50 Matches: 104  
Percent Similarity: 89.17% Conservative: 3  
Best Local Similarity: 86.67% Mismatches: 12  
Query Match: 49.25% Indels: 1  
DB: 10 Gaps: 1  
US-10-021-811-36 (1-206) x US-09-878-574-3147 (1-379)  
QY 1 MetaApLYsLysGIngnInCYsLYsThSerGlnAspProGluValaArgLYsGlyProTTP 20  
DB 23 ATGACAAAGAGCTTGCC--AACAGCTCATGATCTGAAAGTGAAGAGGGGCATGG 79  
QY 21 ThrMetGluGluAspLeuileuemeCAsnTYrilleaIaasnHISglUGluGValTTP 40  
DB 80 ACATATGAAATGATTAATCTTATGATCACTATATATGCCAATACGAGGGAAGGGGTTTG 139  
QY 41 AasnSerleuAlaLYsaIaAlaGlyLeuLYsaArgasnGlyYsserCYsaArgleuArgTTP 60  
DB 140 AACTCTTTGGCAACGCTGCTGAGCTTAACGTAACCGAAAGAGTTGCCGGCTCCGGTGG 199

QY 61 LeuAsnTYrleuArgProaspValaArgArgGlyYasnIleThrProglUGluGlnleuLeu 80  
DB 200 CTAAACTACCTCCGCTGATGTAGAAAGGGAATATTATACCCGAGGAACAGCTTTTG 259  
QY 81 lIemeGluLeuHISAlaLYsTTPGlyYasnArgTTPserYsilleaIaYsHISleuPro 100  
DB 260 ATCATGAACTTCATGCAAGTGGGGAACAGGTGTCCAAATTCGCAAGCATCTACCC 319  
QY 101 GLYArgThAspAsnGlnIleLYsaSnTYrTPArgThArgIleGlnLYsHISleYs 120  
DB 320 GGAAGAGCTGATTAATGAGATTAGACTTACGACGACAGATCCATTAAGCACTCAAG 379  
RESULT 11  
US-09-878-574-50  
/ Sequence 50, Application US/09878574  
/ Patent No. US20020110548A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Byrum, Joseph R.  
/ APPLICANT: La Rosa, Thomas J.  
/ APPLICANT: Thompson, Michael D.  
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
/ TITLE OF INVENTION: Plants  
/ FILE REFERENCE: 38-21 (15401) B  
/ CURRENT APPLICATION NUMBER: US/09/878,574  
/ CURRENT FILING DATE: 2001-12-21  
/ PRIOR APPLICATION NUMBER: 09/333,535  
/ PRIOR FILING DATE: 1999-06-14  
/ NUMBER OF SEQ ID NOS: 15775  
/ SEQ ID NO 50  
/ LENGTH: 350  
/ TYPE: DNA  
/ ORGANISM: Glycine max  
/ FEATURE:  
/ NAME/KEY: unsure  
/ LOCATION: (1) ..(350)  
/ OTHER INFORMATION: unsure at all n locations  
/ OTHER INFORMATION: Clone ID: LIB3028-056-Q1-B1-G4  
US-09-878-574-50  
Alignment Scores:  
Pred. No.: 4.98e-58 Length: 350  
Score: 556.50 Matches: 103  
Percent Similarity: 93.69% Conservative: 1  
Best Local Similarity: 92.79% Mismatches: 6  
Query Match: 48.90% Indels: 1  
DB: 10 Gaps: 1  
US-10-021-811-36 (1-206) x US-09-878-574-50 (1-350)  
QY 1 MetaApLYsLysGIngnInCYsLYsThSerGlnAspProGluValaArgLYsGlyProTTP 20  
DB 16 ATGACAAAGAGCTTGCC--AACAGCTCATGATCTGAAAGTGAAGAGGGGCATGG 72  
QY 21 ThrMetGluGluAspLeuileuemeCAsnTYrilleaIaasnHISglUGluGValTTP 40  
DB 73 ACATATGAAAGAGCTTAATCTTATGATCACTATATATGCCAATACGAGGGAAGGGGTTTG 132  
QY 41 AasnSerleuAlaLYsaIaAlaGlyLeuLYsaArgasnGlyYsserCYsaArgleuArgTTP 60  
DB 133 AACTCTTTGGCAACGCTGCTGAGCTTAACGTAACCGAAAGAGTTGCCGGCTCCGGTGG 192  
QY 61 LeuAsnTYrleuArgProaspValaArgArgGlyYasnIleThrProglUGluGlnleuLeu 80  
DB 193 CTAACTACCTCCGCTGATGTAGAAAGGGAATATTATACCCGAGGAACAGCTTTTG 252  
QY 81 lIemeGluLeuHISAlaLYsTTPGlyYasnArgTTPserYsilleaIaYsHISleuPro 100  
DB 253 ATCATGAACTTCATGCAAGTGGGGAACAGGTGTCCAAATTCGCAAGCATCTACCC 312  
QY 101 GLYArgThAspAenglnIleLYsaSnTYrTPArgThArgIleGlnLYsHISleYs 111  
DB 313 GGAAGAGCTGATTAATGAGATTAAAGAACTACTGG 345

```
RESULT 12
US-09-878-574-194
; Sequence 194, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21 (15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 194
; LENGTH: 365
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-054-Q1-B1-B1
US-09-878-574-194

Alignment Scores:
Pred. No.: 1,1e-56 Length: 365
Score: 545.50 Matches: 102
Percent Similarity: 93.64% Conservative: 1
Best Local Similarity: 92.73% Mismatches: 6
Query Match: 47.93% Indels: 1
DB: 10 Gaps: 1

US-10-021-811-36 (1-206) x US-09-878-574-194 (1-365)

QY 1 MetAspLySlySGInGInCysLysThrSerGlnAerProGluValArgLysGlyProTrr 20
DB 37 ATGGACAAAGAACTTGGC---AACACGCTCATGATCTCTGAAGTGAAGAGGGGCGCATGG 93
QY 21 ThrMetGluGluAspLeuIleuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrr 40
DB 94 ACMAAGAGAGAACTTAAATCTTGAATCCTTATTCGCAATTCACGGGGAAGGGGTTTGG 153
QY 41 AansSerLeuAlaLysAlaAlaGlyLeuLysArgAangLyLysSerCysArgLeuArgTrr 60
DB 154 AACTCTTTGGCCAAAGCGCTGCTGACTTAACGTACCGGAAAGATTGCGCGCTCGGTGG 213
QY 61 LeuAsnTyrLeuArgProAspValArgArgGlyAenIleThrProGluGluLeuLeu 80
DB 214 CTAAACTACCTCCGCTCGATGTGTGAAGAGGGAATATTACCCGGAAGAACAGCTTTTG 273
QY 81 IleMetGluLeuHisAlaLysTrrGlyAsnArgTrrSerLysIleAlaLysHisLeuPro 100
DB 274 ATCAATGGAATCTTCAATGCAAGATGGGGAACAGGTGTGTCCAAATTTGCCAAGCATCTACC 333
QY 101 GlyArgThrAspAsnGluIleLysAsnTyr 110
DB 334 GGAAGGACTGATATGATGATTAAAGACTTAC 363

RESULT 13
US-10-021-811-45
; Sequence 45, Application US/10021811
; Publication No. US20030024007A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Fang, Yiwen
; APPLICANT: Odell, Joan
; APPLICANT: Weng, Zude
; TITLE OF INVENTION: Plant Myb Transcription Factor Homologs
; FILE REFERENCE: B01294 US NA
; CURRENT APPLICATION NUMBER: US/10/021,811
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: 60/110,609
; PRIOR FILING DATE: 1998-December-02
```

```
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 45
; LENGTH: 1236
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (519)
; NAME/KEY: unsure
; LOCATION: (521)
; NAME/KEY: unsure
; LOCATION: (530) . (531)
; NAME/KEY: unsure
; LOCATION: (534)
; NAME/KEY: unsure
; LOCATION: (800)
; NAME/KEY: unsure
; LOCATION: (1124)
; NAME/KEY: unsure
; LOCATION: (1151)
US-10-021-811-45

Alignment Scores:
Pred. No.: 8.25e-54 Length: 1236
Score: 528.00 Matches: 103
Percent Similarity: 68.06% Conservative: 27
Best Local Similarity: 53.93% Mismatches: 43
Query Match: 46.40% Indels: 18
DB: 9 Gaps: 4

US-10-021-811-36 (1-206) x US-10-021-811-45 (1-1236)

QY 1 MetAspLySlyS-----GInGInCysLysThrSerGlnAer 12
DB 96 ATGATGAAAAGAGACAGAACTAGCAACCCCTTTAAAGTTGTAGAGACAGATGAC 155
QY 13 ProGluValArgLysGlyProTrrThrMetGluGluAspLeuIleuMetAsnTyrIle 32
DB 156 -----CTTGAAGAGAGCCCTTGACCGTGATGAAGACCTCACTTATCAATTACAT 209
QY 33 AlaAsnHisGlyGluGlyValTrrAsnSerLeuAlaLysAlaGlyLeuLysArgAan 52
DB 210 GCCACTCATGCGCAAGAGTGTGCGTGAACACGCTGCTCTGCTGGGTGAAGAGCAAG 269
QY 53 GlyLysSerCysArgLeuArgTrrLeuAsnTyrLeuArgProAspValArgArgGlyAan 72
DB 270 GGGAGAGATTGCAAGATTGAGGTGGCTGAATTATCTGCTGCTGATTTGACCGTGAAAC 329
QY 73 IleThrProGluGluGluLeuLeuIleMetGluLeuHisAlaLysTrrGlyAsnArgTrr 92
DB 330 ATCACTTGAAGAAACAACTTTGATTCTGAGACTTCATTTCTGCTGGGGAACCGTTGG 389
QY 93 SerLysIleAlaLysHisLeuProGlyArgThrAspAsnGluIleLysAsnTyrTrrArg 112
DB 390 TCGAAAATTGCTCAATATTTTGGCTGTGAACCGAACAAGATAAAGAACTATTGGAGA 449
QY 113 ThrArgIleGlnLysHisIleLysGlnAlaGlu-----AsnPhcGlnGlnGlnSer 129
DB 450 ACCGGTCTCAAAAGCATGCGCAAGCAACTCAAAATGTGACGTGAATGCAAGCAATTCAG 509
QY 130 SerAsnAsnSerGluIleAsnAspHisGlnAlaSerThrSerHisValSerThrMetAla 149
DB 510 GACACCATGNGNTACCTTTGNNATNCCAAAGCTGTGGAACGATTCACAGACGGCGAC 569
QY 150 GluProMetGluMetLys-----SerProCysTyrGlnGlyMetLeu 164
DB 570 GGGCCCGGTAAACCAACCGGTAACTGGCGCGCCACCAACAAATGCAATTCACCTAACGGRA 629
QY 165 GluProPheSerThrGlnPheProThrIleAsn 175
DB 630 CAACCTTATACCAACCAATTCGAGGTTTGAAT 662
```

## RESULT 14

US-10-021-811-17  
 / Sequence 17, Application US/10021811  
 / Publication NO. US20030024007A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Cahoon, Rebecca E.  
 / APPLICANT: Fang, Yiwen  
 / APPLICANT: Odeh, Joan  
 / APPLICANT: Meng, Zude  
 / TITLE OF INVENTION: Plant Myb Transcription Factor Homologs  
 / FILE REFERENCE: B01294 US NA  
 / CURRENT APPLICATION NUMBER: US/10/021,811  
 / CURRENT FILING DATE: 2001-12-14  
 / PRIOR APPLICATION NUMBER: 60/110,609  
 / PRIOR FILING DATE: 1998-December-02  
 / NUMBER OF SEQ ID NOS: 63  
 / SOFTWARE: Microsoft Office 97  
 / SEQ ID NO: 17  
 / LENGTH: 587  
 / TYPE: RNA  
 / ORGANISM: Oryza sativa  
 / FEATURE:  
 / NAME/KEY: unsure  
 / LOCATION: (577)  
 / NAME/KEY: unsure  
 / LOCATION: (582)..(583)  
 / NAME/KEY: unsure  
 / LOCATION: (587)  
 US-10-021-811-17

## Alignment Scores:

Pred. No.: 2,86e-53 Length: 587  
 Score: 519.50 Matches: 94  
 Percent Similarity: 77.24% Conservative: 18  
 Best Local Similarity: 64.83% Mismatches: 14  
 Query Match: 45.65% Indels: 19  
 Gaps: 2

US-10-021-811-36 (1-206) x US-10-021-811-17, (1-587)

QY 14 GluValArgLysGlyProTTPThrMetGluGluAspLeuLeuMetAsnTyrIleAla 33  
 DB 77 GAGCTCCGCGAGGCGCTGACGCTGAGAGGAGACCTGCTCTCGCACTACATGCGC 136  
 QY 34 AsnHisGlyGluGlyValTTPAsnSerLeuAlaLysAlaGlyLeuLysArgAsnGly 53  
 DB 137 GCCATGCGCGAGGCGCGCTGGAGCGCGCTGCGCGCGCTGGAAGCGGAGCGGG 196  
 QY 54 LysSerCysArgLeuArgTTPLeuAsnTyrLeuArgProAspValArgArgGlyAsnIle 73  
 DB 197 AAGAGCTGCGCGCTGCGCTGCTGACCTGAGCGCGAGCGCTGAGAGGAGGAGAACATG 256  
 QY 74 ThrProGluGluGluLeuLeuIleMetGluLeuHisAlaLysTTPGlyAsnArgTTPSer 93  
 DB 257 ACGGCGAGGAGCGAGCTGATGATCTGAGAGCTCCATGCGCGCTGGGGGAGATCGTGGAGC 316  
 QY 94 LysIleAlaLysHisLeuProGlyArgTTPAspAsnGluIleLysAsnTyrTTPArgTTP 113  
 DB 317 AAGATTCGCGAGCTCTCCCGCGCGAGCGAGCAACAGATCAAGAACTGCGCGAGC 376  
 QY 114 ArgIleGluLysHisIleLysGlnAlaGluAspHseGlnGlnInserSerAsnAsnSer 133  
 DB 377 CCGCTCCAGAGCAAGCGGCAAGCACTC-----AACTGC 409  
 QY 134 GluIleAsnAspHisGln-----AlaSerThrSer 143  
 DB 410 GAGCTCACTCCGAGCAAGCTTCAAGAGCACTATGCGCTACTTGATGCCCGCTCTCG 469  
 QY 144 HisValSerThrMet 148  
 DB 470 AACGATCAACGCTC 484  
 RESULT 15

## US-10-021-811-47

/ Sequence 47, Application US/10021811  
 / Publication NO. US20030024007A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Cahoon, Rebecca E.  
 / APPLICANT: Fang, Yiwen  
 / APPLICANT: Odeh, Joan  
 / APPLICANT: Meng, Zude  
 / TITLE OF INVENTION: Plant Myb Transcription Factor Homologs  
 / FILE REFERENCE: B01294 US NA  
 / CURRENT APPLICATION NUMBER: US/10/021,811  
 / CURRENT FILING DATE: 2001-12-14  
 / PRIOR APPLICATION NUMBER: 60/110,609  
 / PRIOR FILING DATE: 1998-December-02  
 / NUMBER OF SEQ ID NOS: 63  
 / SOFTWARE: Microsoft Office 97  
 / SEQ ID NO: 47  
 / LENGTH: 1181  
 / TYPE: DNA  
 / ORGANISM: glycine max  
 US-10-021-811-47

## Alignment Scores:

Pred. No.: 1,06e-52 Length: 1181  
 Score: 518.50 Matches: 117  
 Percent Similarity: 48.21% Conservative: 31  
 Best Local Similarity: 38.11% Mismatches: 52  
 Query Match: 45.56% Indels: 107  
 Gaps: 9

US-10-021-811-36 (1-206) x US-10-021-811-47 (1-1181)

QY 3 LysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTTPThrMet 22  
 DB 34 AAGAGCTGACGAGCTTCTGAGAGATGACATGAACTTAAAGAGGCGCTGACCTG 93  
 QY 23 GluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTTPAsnSer 42  
 DB 94 GAAAGGAGTAATCTGCTCTCCCAATATATTTTATATATGAGGAGCGGCAATTTG 153  
 QY 43 LeuAlaLysAlaLysGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTTPLeuAsn 62  
 DB 154 CTGGCTTAAACGTTCAAGATTAAGAGAACTGGGAAAGTTGCAAACTTAAGCTGCTAAAT 213  
 QY 63 TyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuIleMet 82  
 DB 214 TATCTAAAGCCAGAGTAAAGCGGGAATTTAACCCCAAGAGCAATTTATTTCTT 273  
 QY 83 GluLeuHisAlaLysTTPGlyAsnArgTTPSerLysIleAlaLysHisLeuProGlyArg 102  
 DB 274 GAACTCCACTCAAAAGTGGGAAACAGGTGTCAAAATTTGCAAACTTTGCCAGGAGCA 333  
 QY 103 ThrAspAsnGluIleLysAsnTyrTTPArgThrArgIleGlnLys-----HisIle 119  
 DB 334 ACAAGCATGAAATCAAGAACTATTGGAGAACTAGATTCAAGAAACAAGCAATTTG 393  
 QY 120 Lys-----GlnAlaGluAspHseGln----- 126  
 DB 394 AAAATTTCACTGACAGAGAGAGTTCAAGAACTTTAGCGCTTTCTGAGTCCCTAGA 453  
 QY 127 -----GlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140  
 DB 454 TTGCTTCAGAAAGCAAAAGATTCATCTTTCAACATGTCATTT-----CAAAACAGGCA 510  
 QY 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProProCysTyr 160  
 DB 511 ATTCTATGCTTTTGAATTATGATTTCTCAGCATTTAATCTTTGGACCAATACCTCTTG 570  
 QY 161 GlnGly-----MetLeuGluProPheSerThr-----GlnPhePro 172  
 DB 571 CAGGAGCTTGTATGAATGAAGAGTGTCCCACTTACATGAGCAACATGAGAGACTCAG 630  
 QY 173 ThrIleAsnProAspGlnSerSerCys----- 182



Db 631 ACTCGAACCACCAATGGTTTCATGCTCTTGTCTGAGTCAGCAAAATATTCAAAA 690  
Qy 183 -----ThrAsnAsp 185  
Db 691 GTGCCTCAGCATTTTGACACACCACCATCCCAATTTCATGCTTGAATACCAATGAC 750  
Qy 186 -----AsnAsnAsn1e-----189  
Db 751 TTTGGCACCCTTCATATGAAAGTTATATGTAACAACAAATGCTATGAGATGACAAAC 810  
Qy 189 -----189  
Db 811 TTCAAAAGACTACTACATGGGTGGCTGAGATGGCAATACCAATTGTGATTGTCAA 870  
Qy 190 -----AsnTyrTyrSerMetGluAsp 196  
Db 871 ATGTAGAGAGCAATTGGGTAAACAACAGCATTTTGATGTATGATGTGAGACATGGATGAA 930  
Qy 197 SerTyrSerMetGlnLeuIeu 203  
Db 931 CTGTGGCAGTTTAGCAGTTA 951

Search completed: February 8, 2003, 12:31:45  
Job time : 62 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 8, 2003, 11:07:20 ; Search time 1928 Seconds

(without alignments)  
1730.431 Million cell updates/sec

Title: US-10-021-811-36  
Perfect score: 1138  
Sequence: 1 MDKXQCKTSGDPEVRKGPW.....NNINYSMEDSWSMQLNGD 206

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=x1h  
-O=/cgm2.1/USPTO.spool/US10021811/runat.03022003.111222.25647/app.query.fasta.1.391  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=prc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-NO XLPXY -NO MAMP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-MAEN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: gb\_gsa:\*  
18: em\_gsa\_hum:\*  
19: em\_gsa\_inv:\*  
20: em\_gsa\_pln:\*  
21: em\_gsa\_vrt:\*  
22: em\_gsa\_fun:\*  
23: em\_gsa\_nam:\*  
24: em\_gsa\_mus:\*  
25: em\_gsa\_other:\*  
26: em\_gsa\_pro:\*  
27: em\_gsa\_rod:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	991	87.1	562	14	BM732121 sal173b11.1
2	859.5	75.5	635	14	BM146360 NF047E06F
3	851	74.8	501	13	BM527606 sal63g06
4	846	74.3	782	10	BE658316 CM700005B
5	816	71.7	552	13	BM527774 sal65g03
6	813.5	71.5	560	14	BM732539 sal178e08
7	808	71.0	669	12	BG457971 NF037A10P
8	796	69.9	682	14	BM146831 NF028G06F
9	795.5	69.9	568	13	BM527664 sal64d10
10	784	68.9	523	9	AI930997 sb45h07.Y
11	784	68.9	688	13	BI272897 NF091A12F
12	779.5	68.5	544	13	BM527508 sal162f08
13	774.5	68.1	542	13	BM528383 sal157f09
14	724	63.6	681	9	AI486576
15	709.5	62.3	476	10	BE057370
16	673.5	59.2	592	10	AW928236 EST307029
17	671	59.0	558	10	BE324639 NF024C04P
18	670.5	58.9	470	10	AW156542 se28d11.Y
19	657	57.7	409	12	BE804790 se45f11.Y
20	654.5	57.5	425	10	AW423958 bh58e05.Y
21	642	56.4	503	14	BQ104458 g91307.e
22	639	56.2	500	9	AI897784
23	639	56.2	519	9	AI897681 EST267124
24	637.5	55.0	502	14	BQ106505 IC0568.e
25	627.5	55.1	447	13	BI273011 NF097F02F
26	621	54.6	395	12	BE805071
27	614	54.0	450	10	BE058947
28	601	52.8	431	10	AW459279
29	595	52.3	586	12	BE820766
30	593	52.1	715	10	BE658737
31	529.5	46.5	542	9	AU238946
32	527	46.3	397	10	AW432364
33	526.5	46.3	727	12	BG607379
34	522.5	45.9	676	14	BU013207
35	514	45.2	766	10	BE659054
36	514	45.2	635	10	AM032656
37	513.5	45.1	690	14	BQ147546
38	513.5	45.1	605	13	BU233398
39	511	44.9	370	12	BG642441
40	510	44.8	371	9	AI487923
41	510	44.6	502	10	BE459188
42	508	44.6	566	13	BI424418
43	506.5	44.5	510	12	BF325282
44	506	44.5	536	13	BI701275
45	501	44.0			

ALIGNMENTS

RESULT 1  
LOCUS BM732121 562 bp mRNA linear EST 01-MAR-2002  
DEFINITION sal173b11.y1 Gm-c1061 glycine max cDNA clone SOYBEAN CLOVE ID:  
Gm-c1061-4678 5' similar to TR:Q39028 Q39028 ATMYB2. ; mRNA  
sequence.

ACCESSION BM732121 GI:19053454  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

soybean.  
Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

REFERENCE  
AUTHORS

1 (bases 1 to 562)  
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schuck, R., Ritters, B., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)

JOURNAL  
COMMENT

Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com  
Seq primer: -40RP from Gldco  
High quality sequence stop: 421.

## FEATURES

## source

Location/Qualifiers

1..562  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-cl061-4678"  
/clone\_id="Gm-cl061"  
/tissue\_type="mature flowers of field grown plants"  
/lab\_host="DH10B"  
/note="vector: pBluescript II SK+; Site 1: EcoRI, Site 2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Radden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 187 a 115 c 128 g 132 t  
ORIGIN

## Alignment Scores:

Pred. No.: 2,7e-106 Length: 562  
Score: 991.00 Matches: 181  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 87.08% Indels: 0  
DB: 14 Gaps: 0

US-10-021-811-36 (1-206) x BM732121 (1-562)

QY 1 MetAspLysIsvGInGInCvLyseThSeGlnAspProGluValAlaGlySGlyProTTP 20  
Db 20 ATGGATTAATAAACAACAGTGAAGCGCTCAAGATCTCTGAAGTGAAGAAAGGGCTTGG 79  
QY 21 ThirMetGluIuAspLeuIleuWeeAsnTyrlleAlaenHISGlyGluGlyValTTP 40  
Db 80 ACAATGGAAGAAGACTTATCTTGATGAATATATTGGAATCATGGGAAGGTTTGG 139  
QY 41 AAmSerLeuAlaIysAlaIaGlyLeuLysArgenGlyIysSerCyBAArgLeuArgTTP 60  
Db 140 AACTCTTTGGCAAGCGTGTGCTCAACGTAACGAAAGAAAGATTCCCGCTAAGGTGG 199  
QY 61 LeuAsnTyrlleuArgProAspValaArgArgIysAsnIleThrProGluGluGlnLeuLeu 80  
Db 200 CTAATATACCTCCGCTCGATGTAGAAAGAGGAATATTACCCGAGGAACAACATTGG 259  
QY 81 I..eMetGluLeuHISAlaIysTyrlleGlyAsnArgTTPSerIysIleAlaIysIleuPro 100

Db 260 ATTATGAGCTCCACGCAAAAGTGGGAAACAGGTGTGTCAAAATTGCCAAGCATCTACT 319

QY 101 GIVArXThAspAsnGluIleLysAsnTyrllePArgThArgIleGlnLysIleIys 120

Db 320 GGAAGACGTGAATGAATGATCAAGAACTATTGGAGACAGGATCCAGAGCAATCAAG 379

QY 121 GlnAlaGluAspNheGlnGlnInSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140

Db 380 CAAAGTGAAGAACTTTTCAGCAACAGAGTAGTAATATTCTGAGATTAATGATCACCAGCT 439

QY 141 SerThSerHisValSerThMetAlaGluProMetGluMetTyrlSerProProCysTyrl 160

Db 440 AGCACTAGCCATGTTTCCACCATGCTGAGCCACGAGATATATTCTCCACCTGTTAT 499

QY 161 GlnGlyMetLeuGluProPheSerThrgInPheProThrIleAsnProAspGlnSerSer 180

Db 500 CAAGGAATGTTAGAGCCATTTCACATCGCTGAGCCACGAGATATATTCTCCACCTGTTAT 559

QY 181 Cys 181

Db 560 TGT 562

RESULT 2

BQ146360 BQ146360 635 bp mRNA linear EST 24-APR-2002

LOCUS NF047E06FL1F050 Developing flower Medicago truncatula cDNA clone

DEFINITION NF047E06FL 5', mRNA sequence.

ACCESSION BQ146360

VERSION BQ146360.1 GI:20283419

KEYWORDS EST

SOURCE barrel medic.

ORGANISM Medicago truncatula

REFERENCE 1 (bases 1 to 635)

AUTHORS Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,

TITLE Flores, H.R., Imman, J.T., Weller, J.W. and May, G.D.

JOURNAL Expressed Sequence Tags from the Samuel Roberts Noble Foundation

COMMENT Medicago truncatula flower library  
Unpublished (2001)  
Contact: May GD  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7391  
Fax: 580 221 7380  
Email: gdmay@noble.org  
Insert Length: 635 Std Error: 0.00  
Plate: 047 row: E column: 06  
Seq primer: TCACACAGAAACAGCTATGAC.

FEATURES  
source location/Qualifiers

1..635  
/organism="Medicago truncatula"  
/db\_xref="taxon:3880"  
/clone="NF047E06FL"  
/clone\_id="NF047E06FL"  
/tissue\_type="Developing flower"  
/tissue\_type="Developing flowers"  
/dev\_stage="Developmentally pooled. Contains a mixture of very young, developing, fully-opened flowers and flowers in early transition into pods."  
/note="vector: Lambda Zap, cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from very young, developing, fully-opened flowers and flowers transitioning into pods. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the GigaPack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using Exaseist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."

BASE COUNT 229 a 121 c 117 g 168 t  
 ALIGNMENT SCORES:  
 Pred. No.: 9,86e-91 Length: 635  
 Score: 859.50 Matches: 167  
 Percent Similarity: 85.29% Conservative: 7  
 Best Local Similarity: 81.86% Mismatches: 18  
 Query Match: 75.53% Indels: 12  
 DB: 14 Gaps: 5

US-10-021-811-36 (1-206) x B0146360 (1-635)

QY 1 MetAspLySylGInGInCySylSFrThSerGlnAspProGluValArgLySglYProTTP 20  
 DB 37 ATGATATAAAAA---CCATGCACCTCATCTCAAGATCCTGAGAGGAAAGGCGCATGG 93

QY 21 ThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGluValTTP 40  
 DB 94 ACCATGGAAGAGAGCTGATTATTAATCAATTATATTCCAATCATGGTGAAAGGTGTTGG 153

QY 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTTP 60  
 DB 154 AATTCCTTAGCCAAAGCTGCTGCTTAAACGTAAGAGAAAGCTTCCAGGCTTCGATGG 213

QY 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGluLeu 80  
 DB 214 TTAACACTACCTCGTCCAGATGTTAGAGAGGGAATATTAACCTGAGAACAACTTTTG 273

QY 81 IleMetGluLeuHisAlaLysTTPGlyAsnArgTTPSerLysIleAlaLysIleuPro 100  
 DB 274 ATCATGGAACCTTCATGGAAGTGGGAAATAGGTGTCACAAATTCGAAAGCATCTTCCA 333

QY 101 GlyArgThrAspAsnGluIleLysAsnTyrTTPArgThrArgIleGluLysHisIleLys 120  
 DB 334 GGAAGAACTGACATGATGATTAAGATTTTGGAGGCTAGATACAAAGCCACATTAAG 333

QY 121 GlnAlaGlu-----AsnPhelGlnGlnInSerSerAsnAsnSerGluIle 135  
 DB 394 CAATGTATCATCTCCACCAAAATATATTCACAAATGAGT-----TTAGAGATA 444

QY 136 AsnAsp---HisGlnAlaSerThrSer-HisValSerThMetAlaGluProMetGluLeu 154  
 DB 445 AATGATCATCATCATCATGAGTAGAGCCAAAGTTTCCAAATGTAGAGCAATGAGATC 504

QY 154 TyrSerProProCysTyrGlnGlyMetLeuGluProPheSerThGlnPheProTTP 174  
 DB 505 TTACTCTCCACCTTCATACCAAGGAACTTGGAGCCATTTCCAGCTCAATTTCCAAACAT 564

QY 174 eAsnProAspGln---SerSerCysCysThrAsnAspAsnAsnIleAsnTyrTTPSe 193  
 DB 565 CAATGATCATCATCTCCCAAGTTGTTGTAACAATGACACAAATACAAATTAATTGGAG 624

QY 193 rMetGluAsp 196  
 DB 625 CATGAGAGAT 634

RESULT 3  
 LOCUS BMS27606 501 bp mRNA linear EST 19-FEB-2002  
 DEFINITION BMS27606.1 Gm-c1061 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1061-3780 5' similar to TR:Q39028 Q39028 ATMYB2.1; mRNA sequence.  
 ACCESSION BMS27606  
 VERSION BMS27606.1 GI:18733434  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseolaeae; Glycine.  
 REFERENCE 1 (bases 1 to 501)

AUTHORS  
 Shoemaker, R., Keim, P., Vodkin, L., Expelding, J., Coryell, V., Khanna, A., Bolla, B., Mair, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Rittner, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R., and Wilson, R.  
 TITLE  
 JOURNAL  
 COMMENT  
 Public Soybean EST Project  
 Unpublished (1999)  
 Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com  
 Seq primer: -40RP from Glibco  
 High quality sequence stop: 421.

FEATURES  
 source  
 1. 501  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="SOYBEAN CLONE ID: Gm-c1061-3780"  
 /clone\_id="Gm-c1061"  
 /issue\_type="mature flowers of field grown plants"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript II SK+; Site 1: EcoRI, Site 2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 170 a 104 c 114 g 113 t  
 ORIGIN

ALIGNMENT SCORES:  
 Pred. No.: 6,89e-90 Length: 501  
 Score: 851.00 Matches: 157  
 Percent Similarity: 96.34% Conservative: 1  
 Best Local Similarity: 95.73% Mismatches: 6  
 Query Match: 74.78% Indels: 0  
 DB: 13 Gaps: 0

US-10-021-811-36 (1-206) x BMS27606 (1-501)

QY 1 MetAspLySylGInGInCySylSFrThSerGlnAspProGluValArgLySglYProTTP 20  
 DB 10 ATGATATAAAAACTGTCACACGTCCTCAAGTCTCGAAGTAGAAGAGACCTTGG 69

QY 21 ThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGluValTTP 40  
 DB 70 AGCATGGAAGAGAGCTGATTGATCTGATCAATATATTCGAAATCATGGGAAAGGTGTTGG 129

QY 41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTTP 60  
 DB 130 AATTCCTTAGCCAAAGCTGCTGCTTCAAGCTACCGAAGAGTTCGCGCTAAGTGG 189

QY 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGluLeu 80  
 DB 190 CTAAACTACCTCGCTCGATGTTAGAGAGGAAATATTAACCCGAGAACAACTTTG 249

QY 81 IleMetGluLeuHisAlaLysTTPGlyAsnArgTTPSerLysIleAlaLysIleuPro 100  
 DB 250 ATCATGAGCTTCACGCAAGTGGGAAAGGTGTCACAAATTCGCAAGCATCTACT 309

```

Qy 101 GYAGTThrAspAsnGluIleLysAsnTYrTPArGThrArgIleGlnLYSHieLys 120
Db 310 GGTAGACAGATTAATGATCATCAAGAACTATTGGAGGACCAAGATCCAGACATCAAG 369
Qy 121 GlnAlaGluAnPheGlnGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db 370 CAAGCTGGAACCTTCAACAACAATTGATTAACCTTGAGATTAAGATCAACCAAGCT 429
Qy 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTYrSerProCysTYr 160
Db 430 AACAATGACCAATGTTTACCATGCGTCAACCAATGAGACCACTATTCCACCCCTTTAT 489
Qy 161 GlnGlyMetLeu 164
Db 490 CAAGGAATGTGA 501

RESULT 4
BE658316/c
LOCUS BE658316
DEFINITION mRNA sequence.
ACCESSION BE658316
VERSION BE658316
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseleae;
Glycine
1 (bases 1 to 782)
Vodkin,L., Keim,P., Shoemaker,R., Retzel,E., Khanna,A., Corryell,V.,
Expelling,J., Rapph,C., Shoop,B., Pardinas,J., Liu,L. and Lewin,H.
A Functional Genomics Program for Soybean (NSF 9872565)
Unpublished (1999)
Other ESTs: AI930997 corresponding to Gm-c1015-278 (5')
Contact: Vodkin, L.O., PI, A Functional Genomics Program for
Soybean (NSF 9872565)
Lewin, H. A., Director, Keck Center for Comparative and Functional
Genomics
University of Illinois
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
Tel: (217) 244-6147
Fax: (217) 333-4582
Email: l-vodkin@uiuc.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134. For further information
call: (800) 430-0030 or (314) 427-3322 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or info@genome
systems.com web site: www.genomesystems.com
Seq primer: 5'-TTTTTTTTTTTTTTTTTTT(A/C/G)-3'.
Location/Qualifiers
1..782
/organism="Glycine max"
/db_xref="taxon:3847"
/cclone="Gm-r1070-1759"
/cclone_lib="Gm-r1070"
/note="The library Gm-r1070 is a sequence-driven, rerecked
set of 9,216 clones selected from cDNA libraries from
various tissues and stages of development of soybean that
represent 2,639 sequences from immature cotyledons, 1,770
from immature seed coats, 3,938 from flowers, and 869
from young pods. The 5' ESTs of the source clones from
the different libraries was used to select singletons, or
a representative of each contig, which were rerecked to
form library Gm-r1070. The cDNA clones of the rerecked
Gm-r1070 library were then sequenced at the 3' end. The
contig analysis to select unique genes was performed by
the laboratory of Ernest Retzel, Center for Computational
Genomics and Bioinformatics, University of Minnesota,
http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html
. Rerecking was performed by Genome Systems, St. Louis,

```

http://www.genomesystems.com, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, http://www.life.uiuc.edu/biotech/keck.html. Note: The corresponding 5' EST from each clone in the Gm-r1070 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Genome Systems clone id of the original cDNA library that is also listed under 'OTHER EST'."

```

BASE COUNT 215 a 144 c 143 g 253 t 27 others
ORIGIN

Alignment Scores:
Pred. No.: 5,256-89 Length: 782
Score: 846.00 Matches: 162
Percent Similarity: 90.61% Conservative: 2
Best Local Similarity: 89.50% Mismatches: 14
Query Match: 74.34% Indels: 3
DB: 10 Gaps: 2

US-10-021-811-36 (1-206) x BE658316 (1-782)

Qy 29 MetAsnTYrIleAlaAsnHisGlyGluGlyValTPArnSerLeuAlaLYSHieLys 48
Db 780 ATCAACTATATGTCACAAATCATNNNNNGGTGTTGNANNNCVTGGCCAAAGCTGCTNNT 721
Qy 49 LeuLYSArgAsnGlyLYSserCysArgLeuArgTTPLeuAsnTYrLeuAlaYrProAspVal 68
Db 720 CTCAAACGTAACCGGAAGAAGGTGGCCGGCTAAGTGCTTAACCTACCTCCGCTCTGATGTT 661
Qy 69 ArgArgGlyAsnIleThrProGlnGluGlnLeuIleMetGluLeuHisAlaLYSTTP 88
Db 660 AGAAGAGGGAATATNNNACCAGGAACAACCTTGATCATGAGACTTCACGCAAAAGTGG 601
Qy 89 GLYAsnArgTTPSerLYSileAlaLYSHisLeuProGlyArgTYrThrAspAsnGluIleLYS 108
Db 600 GGAACAAGGTGTGCCAAATTCACCAAGCATCTACCTGGTAGGACAGATTAAGATGAAG 541
Qy 109 AsnTYrTPArGThrArgIleGlnLYSHisIleLYSGlnAlaGluAnPheGlnGlnGln 128
Db 540 AACTATTGGAGAGCCAGGATCCAGAACACATCAAGCAAGCTGAAACCTTTCAGCAACA 481
Qy 129 SerSerAsnAsnSerGluIleAsnAspHisGlnAlaSerThrSerHisValSerThrMet 148
Db 480 ATCAGCAATTAACCTGAGATTAATGATCACCAAGCTACACATGATTTCTTACCATG 421
Qy 149 AlaGluProMetGluMetTYrSerProCysTYrGlnGlyMetLeuGluProPheSer 168
Db 420 GCTGAACCATGAGACCTATTCTCACCCCTTTTATCAAGATGTTAGAGCATTTTCT 361
Qy 169 Thr--GlnPheProThrIleAsnProAspGlnSerSerCysTYrThrAsnAspAsn 187
Db 360 TCAATTCAGTTCCCAACATTAATCTGATCAATCAGTTGTTGACCAATGACCAAC 301
Qy 188 Asn--IleAsnTYrTPSerMetGluAspSerTPSerMetGln-LeuLYSArgIleLYS 206
Db 300 AACAGATTAACCTATTGGAGCATGAGATATCTGCTCAATGATAGTACTGAAACGGGGA 241
Qy 206 P 206
Db 240 T 240

RESULT 5
BM527774
LOCUS BM527774
DEFINITION BM527774 552 bp mRNA linear EST 19-FEB-2002
Gm-c1061-4134 5' similar to TR:049020 O49020 MYB-LIKE DNA-BINDING
DOMAIN PROTEIN. ;, mRNA sequence.
ACCESSION BM527774
VERSION BM527774.1 GI:18733722
KEYWORDS EST.
SOURCE soybean.

```

ORGANISM Glycine max  
Bikaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

REFERENCE 1 (bases 1 to 552)  
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna,  
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers,  
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk,  
R., Riter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann,  
R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact: ccu@resgen.com web site:  
www.resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 426.  
Location/Qualifiers

FEATURES  
source  
1..552  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-c1061-4134"  
/clone\_1lb="Gm-c1061"  
/tissue\_type="mature flowers of field grown plants"  
/lab\_host="DH10B"  
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from mature flowers of field grown plants for the cultivar  
Raiden. Complementary DNA was synthesized from mRNA using  
a primer consisting of a poly(dT) sequence with a XhoI  
restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the pBluescript vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (GibcoBRL). This library was constructed in the  
laboratory of Dr. Randy Shoemaker."

BASE COUNT 184 a 118 c 111 g 139 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.05e-85 Length: 552  
Score: 816.00 Matches: 152  
Percent Similarity: 95.62% Conservative: 1  
Best Local Similarity: 95.00% Mismatches: 5  
Query Match: 71.70% Indels: 2  
DB: 13 Gaps: 2

US-10-021-811-36 (1-206) x BM527774 (1-552)

Oy 49 leu1ysa7ga8ng1y1ys8serCyaa7g1eua7gt7p1eua7nt7y7leua7p7ro7asp7ya1 68  
Db 1 CTCAAAACGTAACCGAAAGAGTTGCCGCTAAAGGTGGCTAAACTACCTCCGTCGTATGTT 60  
Oy 69 A7GA7GG1A7n7le7Th7p7ro7g1n7g1n7leu7n1e7me7G1u7leu7H1a1a7y7trp 88  
Db 61 A7AA7AGG7AA7TA7TA7CA7CCG7A7GA7AA7CA7CTT7TA7G7AT7CA7G7A7CTT7CA7G7AA7G7TG 120  
Oy 89 G1A7n7A7n7g7r7p7se7r7y1e7a1a7y7sh1s7leu7p7ro7G1A7r7Th7A7s7p7a7n7g1u7le7y7s 108  
Db 121 GGA7AA7CA7G7T7G7T7CA7AA7ATT7G7CA7A7G7AT7CA7CT7G7T7A7G7A7CA7G7AT7A7G7T7CA7G 180  
Oy 109 Aa7n7y7T7r7p7A7r7Th7A7r7g1e7g1n7y7sh1s7le7y7G1n7A1a7G1u7a7n7p7he7g1n7g1n7 128

Db 181 AACTATTGAGAGCAGCATCCAGAGCAGCATCAAGAGCTGAGAACTTTCAGCAACAA 240  
Oy 129 Ser7Ser7Aa7n7Se7r7G1u7le7Aa7n7sh1s7G1n7A1a7Se7r7Th7r7Se7H1a7Se7Th7r7me7t 148  
Db 241 ATT7A7TA7TA7TA7CT7GA7G7AT7A7TA7AT7AT7CA7CC7A7G7AT7CA7G7AT7CA7G7AT7T7T7CA7CC7A7G 300  
Oy 149 A1a7G1u7p7ro7me7G1u7me7Ty7se7r7p7ro7Cy7r7y7g1n7g1y7me7le7u7p7ro7p7he7se7r 168  
Db 301 GCT7GA7A7CC7AT7GA7A7G7A7CT7AT7T7CC7A7CC7T7T7AT7CA7G7A7AT7T7A7G7A7CC7A7T7T7CT 360  
Oy 169 Thr---Gln7p7ro7Th7r7le7Aa7n7p7ro7p7g1n7se7r7Cy7e7Th7r7A7s7p7a7n7 187  
Db 361 TCA7TT7CA7G7T7T7CC7CA7CA7TT7AT7AT7CT7TA7T7CA7T7CA7G7T7G7T7GA7CA7A7G7A7CA7A7C 420  
Oy 188 Aa7n---1le7Aa7n7y7T7r7p7Se7r7Me7G1u7a7p7Se7r7p7Se7r7Me7G1n7le7u7Aa7n7G1y7a7p 206  
Db 421 AAC7A7G7AT7TA7CT7AT7T7G7A7G7AT7G7A7G7AT7AT7CT7G7T7CA7T7CA7CT7T7A7G7A7CG7GG7AT 480

RESULT 6  
BM732539  
LOCUS  
DEFINITION  
BMT32539 560 bp mRNA linear EST 01-MAR-2002  
sal78e08.y1 Gm-c1061 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-c1061-5223 5' similar to TR:Q39028 Q39028 ATMBD2.; mRNA  
sequence.

ACCESSION  
BM732539  
VERSION  
BM732539.1 GI:19053872  
KEYWORDS  
EST.

SOURCE  
soybean.  
ORGANISM  
Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

REFERENCE 1 (bases 1 to 560)  
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna,  
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers,  
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk,  
R., Riter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann,  
R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact: ccu@resgen.com web site:  
www.resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 430.  
Location/Qualifiers

FEATURES  
source  
1..560  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-c1061-5223"  
/clone\_1lb="Gm-c1061"  
/tissue\_type="mature flowers of field grown plants"  
/lab\_host="DH10B"  
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from mature flowers of field grown plants for the cultivar  
Raiden. Complementary DNA was synthesized from mRNA using  
a primer consisting of a poly(dT) sequence with a XhoI  
restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the pBluescript vector. The

ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 191 a 128 c 111 g 130 t

ORIGIN

Alignment Scores:

Pred. No.:	2.12e-85	Length:	560
Score:	813.50	Matches:	154
Percent Similarity:	85.71%	Conservative:	8
Best Local Similarity:	81.48%	Mismatches:	22
Query Match:	71.49%	Indels:	5
DB:	14	Gaps:	3

US-10-021-811-36 (1-206) x BM732539 (1-560)

```

QY 9 ThrSerGlnAppProGluValArgLysGlyProTThrMetGluGluAspLeuIleu 28
   :::::
Db 3 TCATCTCATGATCCGAAAGTGAGAAAGGACCATGACCATGAAAGACTTATCTTG 62
   :::::
QY 29 MetAenTyrlleAlaAsnHisGlyGluGlyValTTPAenSerLeuAlaLysAlaGly 48
   :::::
Db 63 ATAACTATATTCGAAATACCGTGAAGTGTGGAACTCTTACCAAGCTTCTGCT 122
   :::::
QY 49 LeuLysAgaAngLysSerCysArgLeuArgTTPleuAenTyrlleuArgProAspVal 68
   :::::
Db 123 CTCACACGACGGGAAAGAGTGTGACTCCGTGGCTTAACTACCTTCGCTGATGTT 182
   :::::
QY 69 ArgArgGlyAenIleThrProGluGluLeuLeuIleMetGluLeuHisAlaLysTrp 88
   :::::
Db 183 AGAAGAGAAATTTTATACCCGAGAACAGCTTTGATCATAGAACTTCATGAAAGTGG 242
   :::::
QY 89 GlyAenAgtTTPSerLysIleAlaLysHisLeuProGlyArgTTPAenGluIleLys 108
   :::::
Db 243 GCGAATAGTGTGCTCAAAATTCGAAAGCATCTTCCAGAAACCTGACAAATGAGATTAG 302
   :::::
QY 109 AsnTyrlTPArgThrArgIleGluLysHisIleLysGluAlaGluAenPheGlnGln 128
   :::::
Db 303 AACTTCTGAGAACAGGATCCAAAGACATTAAGCAAGCTTGACCTTCAACAACAT 362
   :::::
QY 129 SerSerAenSerGluIleAsnAspHisGlnAla-----SerThrSerHisValSer 146
   :::::
Db 363 -----GGTATTCAGAGAAATATGATCATCAACAGCAACATTAATCAAGATGCC 416
   :::::
QY 147 ThrMetIleGluPrometGluMetLysSerProProCysTyrlGlnIleMetLeuGlnPro 166
   :::::
Db 417 ACATGGACATCCAAATGAGACTTTCCTCCACCTCATACCAAGAACCTTTTGACCCA 476
   :::::
QY 167 PheSerThrGlnPheProThrIleAsnProAspGlnSerSerCysCysThrAenAspAsn 186
   :::::
Db 477 TTCCAACCTCATATTCCTACATCACT---GATCAATCAAGTTGTGTACCAACCAAC 533
   :::::
QY 187 AsnAenIleAenTyrlTPSerMetGlu 195
   :::::
Db 534 GACAAACAACTATTTGAGCATCGAG 560
   :::::

```

RESULT 7

BG457971

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

barrel medic.

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;

Medicago.

1 (bases 1 to 669)

REFERENCE

AUTHORS

Liu,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores

H.R., Inman,U.T., Weller,U.W., May,G.D. and Harrison,M.J.

TITLE  
Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
Medicago truncatula phosphate-starved leaf library  
Unpublished (2000)

JOURNAL

COMMENT

Contact: Harrison MJ  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7325  
Fax: 580 221 7380  
Email: mjharrison@noble.org  
Insert Length: 669 Std Error: 0.00  
Plate: 037 row: A column: 10  
Seq primer: TCACACAGAAACACTATGAC.  
Location/Qualifiers

FEATURES

source

1..669

/organism="Medicago truncatula"

/db\_xref="taxon:3880"

/clone="NF037A10PL"

/clone\_lib="phosphate starved leaf"

/tissue\_type="leaf"

/dev\_stage="trifoliolate"

/note="Vector: Lambda Zap; At the trifoliolate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this 30 day period, the plants were fertilized twice weekly with 1/2 Hoaglands solution containing only 20mM potassium phosphate. RNA was prepared from above ground tissues."

BASE COUNT 225 a 141 c 127 g 176 t

ORIGIN

Alignment Scores:

Pred. No.:	1.23e-84	Length:	669
Score:	808.00 <td>Matches:</td> <td>159 </td>	Matches:	159
Percent Similarity:	81.86% <td>Conservative:</td> <td>17 </td>	Conservative:	17
Best Local Similarity:	73.95% <td>Mismatches:</td> <td>19 </td>	Mismatches:	19
Query Match:	71.00% <td>Indels:</td> <td>20 </td>	Indels:	20
DB:	12	Gaps:	7

US-10-021-811-36 (1-206) x BG457971 (1-669)

```

QY 1 MetAenTyrlleuAlaLysGlnCysLysThrSerGlnAppProGluValArgLysGlyProTTrp 20
   :::::
Db 50 ATGACAAAGAAAA---GAGTGCAAGTCTTCCACAAAGATCTGTATGACAAAGGACCATGG 106
   :::::
QY 21 ThrMetGluGluAspLeuIleLeuMetAenTyrlleAlaAsnHisGlyGluGlyValTTP 40
   :::::
Db 107 ACAATGAAAGAAAGCTTGATCTTATCAACTATATATTCATCATGATGGAAGGTGTTGG 166
   :::::
QY 41 AenSerLeuAlaLysAlaLysGlnIleuLysArgAsnGlyLysSerCysArgLeuArgTTP 60
   :::::
Db 167 AACTCTTGGCTAAATCTGCTGCTGTAAACGTACCGGAAAGAGTTGACGGCTTCGGTGG 226
   :::::
QY 61 LeuAenTyrlleuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80
   :::::
Db 227 CTAAACTATCTTCGACCGGAGTGTAGACGAGGAAATTTACACCTGAGAACCACTCTTG 286
   :::::
QY 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTTPSerLysIleAlaLysHisLeuPro 100
   :::::
Db 287 ATCATTTGAATCTTCATGCTCAAGTGGGAAACGCTGCTGAAAATTGCGAACAATCTACCG 346
   :::::
QY 101 GlyArgThrAspAenGluIleLysAenTyrlTPArgThrArgIleGluLysHisIleLys 120
   :::::
Db 347 GGAAGAACCGATTAATGAGATTAAGAACTACGAGGACCAAGAGTCCAAAAGCATATCAAG 406
   :::::
QY 121 GlnAlaGluAenPheGln---GlnGlnSerSerAenSerGluIleAenAspHisGln 139
   :::::
Db 407 CAAGCTGGAAGACTTCATCTCAAGAACAGGTTCTGATATTCAAATTAATGATGATTAAT 466
   :::::
QY 140 -----AlaSerThrSerHisValSerThrMetAlaGluPrometGlu 153
   :::::
Db 467 AATATAATCAACTTATAGGAAGCAAGCCAAATTTCCAACTTGCTGAACCTATGAC 526
   :::::
QY 154 -----MetTyrlSerProProCysTyrlGlnIleMetLeuGluProPheSerThr----- 169
   :::::

```



```

Db      527 ACAATATTATATACACCTCTTACCAGAAATTTTGAACCATTTCCACCTCATCAT 586
Qy      170 GlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsn189
Db      587 CAGTTCCTCCATCATCTCT---GATCAATCAGGTTGTTCT-----622
Qy      190 AsnTyTrpSerMetGluAspSerTrpSerMetGlnLeuAsn 204
Db      623 AACTACTGAGCATGGAAGATCTCTGTCTCACTTCACTCACTTAAT 667

RESULT 8
LOCUS   B0146831 682 bp mRNA linear EST 24-APR-2002
DEFINITION NF028G06FL1J051 Developing flower Medicago truncatula cDNA clone
ACCESSION B0146831
VERSION   B0146831.1 GI:20283890
KEYWORDS EST.
SOURCE   barrel medic.
ORGANISM Medicago truncatula
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
          Medicago.
          1 (bases 1 to 682)
          Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
          Flores,H.R., Imman,J.T., Weller,J.W. and May,G.D.
          Expressed Sequence Tags from the Samuel Roberts Noble Foundation
          Medicago truncatula flower library
          Unpublished (2001)
          Contact: May GD
          Plant Biology Division
          The Samuel Roberts Noble Foundation
          2510 Sam Noble Parkway, Ardmore, OK 73402, USA
          Tel: 580 221 7391
          Fax: 580 221 7380
          Email: gdmay@noble.org
          Insert length: 682 Std Error: 0.00
          Plate: 028 row: G column: 06
          Seq primer: TCACACGAGAAACAGCTATGAC.
          Location/Qualifiers
            1..682
            /organism="Medicago truncatula"
            /db_xref="taxon:3880"
            /clone="NF028G06FL"
            /clone_1lb="Developing flower"
            /tissue_type="Developing flowers"
            /dev_stage="Developmentally pooled. Contains a mixture of
            very young, developing, fully-opened flowers and flowers
            in early transition into pods."
            /note="Vector: Lambda Zap; cDNA was prepared from polyA-
            enriched, pooled samples of equivalent amounts of total
            RNA from very young, developing, fully-opened flowers and
            flowers transitioning into pods. The cDNA was
            directionally ligated into the Uni-Zap XR vector
            (Stratagene) and packaged using the GigaPack III Gold
            packaging extracts. Phagemids containing cDNA inserts were
            in vivo excised from the recombinant Uni-Zap XR vector
            using Exsist helper phage and the E. coli strain
            XL1-Blue MRF' (Stratagene). Excised plasmids were plated
            using SOLR cells."

BASE COUNT 230 a 142 c 128 g 180 t 2 others
ORIGIN
Alignment Scores: 3.29e-83 Length: 682
Score: 796.00 Matches: 160
Percent Similarity: 81.19% Conservative: 17
Best Local Similarity: 73.39% Mismatches: 20
Query Match: 69.95% Indels: 21
DB: 14 Gaps: 7

```

```

US-10-021-811-36 (1-206) x B0146831 (1-682)
Qy      1 MetAspLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
Db      47 ATGGACAAATAAA---GAGTGCAGTTCCTTACAAAGATCTGTATGACGAAGGACCATGG 103
Qy      21 ThrMetGluGluAspLeuLleuMetAsnTyTrlIleAlaAsnHisGlyGluGlyValTrp 40
Db      104 ACAATGGAAGAAAGACTTATCTTGATTCATCAATATATTCGCAATCTGGGAAGGGTTGG 163
Qy      41 AsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuAspTrp 60
Db      164 AACTCTTGCTGTAATTCCTGCTGCTTCTTAACCTACCGAAGAGATTGACGCTTCGGTGG 223
Qy      61 LeuAsnTyTrpLeuArgProAspValArgArgLysAsnIleThrProGluGlnLeuLeu 80
Db      224 CTTAACTATCTTCGACCGGATGTTAGACGAGGAATATTACACCTGAGAGAACACTCTTG 283
Qy      81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
Db      284 ATCATTTGAACCTTCATGCTTAAGTGGGAAACAGTGGTGGAAATTCGCAACATCTACCG 343
Qy      101 GlyArgThrAspAsnGluLleLysAsnTyTrpArgThrArgIleGlnLysHisLys 120
Db      344 GGAAGAACCGAATAAGATAAGATAAGAACTACGAGGACACAGGATCCAAAGCATATTCAG 403
Qy      121 GlnAlaGluAsnPheGln---GlnGlnSerSerAsnAsnSerGluIleAsnAspHisGln 139
Db      404 CAAGCTGAGAACTCTCAATCTCAATCAAGACAGTCTGTATTTCAATAAATGAATGATAAAT 463
Qy      140 -----AlaSerThrSerHisValSerThrMetAlaGluProMetGln 153
Db      464 AATTAATTAATCACTTATATGAGGAACACAGCAACCAATTTCCAACTGCTGAACCTATGGA 523
Qy      153 u-----MetTySerProProCysTyArgGlnGlyMetLeuGluProPheSerThr----- 169
Db      524 ACAATATTATATATACACACCTCTTACCAGAAATTTTGGAAACCATTTCCACCTCATCA 583
Qy      170 -GlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsn11 189
Db      584 TCAGTTCCTCCATCATCTCT---GATCAATCA-----GGTTGATC 619
Qy      189 eAsnTyTrpSerMetGluAspSerTrpSerMetGlnLeuLeuAsnGlyAsp 206
Db      620 TAACACTGAGACATGGAAGATCTCTGTGCACTTCACTCACTTAATGAGAT 671

RESULT 9
LOCUS   BM527664 568 bp mRNA linear EST 19-FEB-2002
DEFINITION sal64d10.y1 Gm-c1061 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-c1061-4003 5' similar to TR:Q39028 Q39028 ATMYB2. ;, mRNA
sequence.
ACCESSION BM527664
VERSION   BM527664.1 GI:18733532
KEYWORDS EST.
SOURCE   soybean.
ORGANISM Glycine max
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
          Glycine.
          1 (bases 1 to 568)
          Shoemaker,R., Keim,P., Vodkin,L., Erpelidg,J., Coryell,V., Khanna
          ,A., Bolla,B., Marrs,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
          Wylie,T., Underwood,K., Stepcoe,M., Theising,B., Allen,M., Bowers
          ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
          ,R., Ritzer,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
          ,R., Waterston,R. and Wilson,R.
          Public Soybean EST Project
          Unpublished (1999)
          Contact: Shoemaker R/Public Soybean EST Project
          Washington University School of Medicine

```

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntville, AL 35801 For further information  
 call: (800)-533-4363 or contact: ccu@resgen.com web site:  
 www.resgen.com  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 422.

## FEATURES

## SOURCE

1 568  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="SOYBEAN CLONE ID: Gm-c1061-4003"  
 /clone\_id="Gm-c1061"  
 /tissue\_type="mature flowers of field grown plants"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
 XhoI; The cDNA library was constructed from mRNA isolated  
 from mature flowers of field grown plants for the cultivar  
 Ralden. Complementary DNA was synthesized from mRNA using  
 a primer consisting of a poly(dT) sequence with a XhoI  
 restriction site. EcoRI adapters were ligated to the  
 blunt-ended cDNA fragments followed by XhoI digestion. The  
 cDNA fragments were directionally cloned into the  
 EcoRI-XhoI restriction site of the pBluescript vector. The  
 ligated cDNA fragments were transformed into DH10B host  
 cells (GibcoBRL). This library was constructed in the  
 laboratory of Dr. Randy Shoemaker."

## BASE COUNT

174 a 134 c 124 g 136 t

## ALIGNMENT SCORES:

Pred. No.: 2.85e-83 Length: 568  
 Score: 795.50 Matches: 153  
 Percent Similarity: 87.91% Conservative: 7  
 Best Local Similarity: 84.07% Mismatches: 19  
 Query Match: 69.90% Indels: 3  
 DB: 13 Gaps: 3

US-10-021-811-36 (1-206) x BM527664 (1-568)

QY 1 MetAspLysLysGlnGlnCysLysThrSerGlnAspProGlnValArgLysGlyProTyr 20  
 Db 29 AUGGACGAAAGCTTGCG--AACACGCTCATGATCTCTGAAGAGAAAGGGCCATGG 85  
 QY 21 ThrMetGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTyr 40  
 Db 86 ACAAAGGAAAGACTTAATCTTGAATCACTATATTCGCAATCAAGGGAAGGGGTTGG 145  
 QY 41 AsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTyr 60  
 Db 146 AACCTTTGGCGAAGGCGCTGAGACTTAACGTAACCGAAGAGATTGCCGCTCCGGTGG 205  
 QY 61 LeuAsnTyrLeuArgProAspValArgArgLysAsnIleThrProGluGluGlnLeu 80  
 Db 206 CTAAACCTACCTCCGCTCGATGTTAGAGAGGAAATATTAACCCGAGAACACCTTTTG 265  
 QY 81 IleMetGluLeuHisAlaLysTyrGlyAsnArgTyrSerLysIleAlaLysHisLeuPro 100  
 Db 266 ATCATGAACTTCATGCAAGAGTGGGAAACAGGTGCTCAAAATTTGCCAAGCATCTACC 325  
 QY 101 GlyArgThrAspAsnGluIleLysAsnTyrTTPArgThrArgGlnGlnLysHisIleLys 120  
 Db 326 GAAAGGACTGATTAATGAATTAAGAACTACGAGGAGCAAGAGATCCAAAGCACTTCAG 385  
 QY 121 GlnAla--GluAsnPhgGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGln 139  
 Db 386 CAAAGCTTCACAGAGCTTCACAGCAAGAGTAGTATCTGAGATTAATTTATCATCCCAA 445  
 QY 140 AAspThrSerHisValSerThrMetAlaGluProMetGluMetCysTyrSerProProCys 159

Db 446 GCTTGACCTAGCAAGATGTCACACATGCGGACGCCCATAGAAACCTATTTCACCCACT 505  
 QY 160 TyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSer 179  
 Db 506 TATCAAGAGATGTTAGATCATTTTCATTCAGTCCCAACA---AATCCATCATCTCT 562  
 QY 180 SerCys 181  
 Db 563 AGTTGT 566

## RESULT 10

AI930997 523 bp mRNA linear EST 30-NOV-2001  
 LOCUS sb45h07.y1 Gm-c1015 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 DEFINITION Gm-c1015-278 5' similar to TR:Q39028 Q39028 ATMYB2.1, mRNA  
 sequence.

## ACCESSION

AI930997  
 VERSION AI930997.1 GI:566961

## KEYWORDS

EST.

## SOURCE

soybean.

## ORGANISM

Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

## REFERENCE

1 (bases 1 to 523)

## AUTHORS

Shoemaker R., Keim, P., Vodka, L., Erpelting, J., Coryell, V., Khanna,  
 A., Bolla, B., Maria, M., Hillier, L., Kudava, T., Martin, D., Beck, C.,  
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers,  
 Y., Peterson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk,  
 R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann,  
 R., Waterston, R. and Wilson, R.

## TITLE

Public Soybean EST Project

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Shoemaker R./Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: ccu@resgen.com  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 422.

## FEATURES

## SOURCE

1 523  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1015-278"  
 /clone\_id="Gm-c1015"  
 /tissue\_type="mature flowers, field grown plants"  
 /lab\_host="XL10-Gold"

## NOTE

/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:  
 XhoI; This cDNA library was constructed from mRNA isolated  
 from mature flowers of field grown plants. The cDNA  
 library was prepared using the Stratagene pBluescript II  
 XR cDNA library construction kit. Complementary DNA was  
 synthesized from mRNA using a primer consisting of a poly  
 (dT) sequence with a blunt restriction site. EcoRI adapters  
 were ligated to the blunt-ended cDNA fragments followed by  
 XhoI digestion. The cDNA fragments were directionally  
 cloned into the EcoRI-XhoI restriction site of the  
 pBluescript vector. The ligated cDNA fragments were  
 transformed into XL10-Gold host cells. This library was  
 constructed by Dr. Randy Shoemaker and Dr. John  
 Erpelting."

## BASE COUNT

172 a 124 c 111 g 115 t 1 others

## ORIGIN

Alignment Scores: 5.68e-82 Length: 523  
 Pred. No.:



Db 622 JACTAGCAGCAGTGAATCTTGTGACCTTCA-CTACTAATGANGAT 671

## RESULT 12

BMS27508

LOCUS BMS27508 544 bp mRNA linear EST 19-FEB-2002  
sa162ef08.y1 Gm-cl061 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-cl061-3735 5' similar to TR:Q39028 Q39028 ATMYB2. ;, mRNA  
sequence.

ACCESSION BMS27508  
VERSION BMS27508.1 GI:18733261  
KEYWORDS EST.

SOURCE soybean.  
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 544)  
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna  
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers  
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk  
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann  
R., Waterston,R. and Wilson,R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800)-533-4363 or contact: ccu@resgen.com web site:  
www.resgen.com

TITLE JOURNAL  
COMMENT

Seq primer: -40RP from Glibco  
High quality sequence stop: 422.  
Location/Qualifiers  
1. 544  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-cl061-3735"  
/clone\_id="Gm-cl061"  
/tissue\_type="mature flowers of field grown plants"  
/lab\_host="DH10B"  
/note="Vector: Bluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from mature flowers of field grown plants for the cultivar  
Raiden. Complementary DNA was synthesized from mRNA using  
a primer consisting of a poly(dT) sequence with a XhoI  
restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the Bluescript vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (GibcoBRL). This library was constructed in the  
laboratory of Dr. Randy Shoemaker."

BASE COUNT 172 a 129 c 122 g 121 t

## ALIGNMENT SCORES:

Pred. No.: 2.04e-81 Length: 544  
Score: 779.50 Matches: 150  
Percent Similarity: 88.70% Conservative: 7  
Best Local Similarity: 84.75% Mismatches: 17  
Query Match: 68.50% Indels: 3  
DB: 13 Gaps: 3

US-10-021-811-36 (1-206) x BMS27508 (1-544)

QY 1 MetAspIyGlnGlnGlnCysIySThrSerGlaAspProGluValArgIySgIyProTyr 20

Db 14 ATGACAAAGAGCTTGGC---AACAGTCTCATGATCCGAAAGTAGAAAGGGCCATGG 70

QY 21 ThmetGluGluAspLeuIleLeuMetAenTyrIleAAsnHsGlyGluGlyTyr 40

Db 71 ACAATGGAAGAGACTTATATCTTGATCACCCTATATTGCCATATCCAGGGGAAGGGTTGG 130

QY 41 AAsnSerIleuAlaIysAlaIagIyLeuIySArgAAsnGlyIySerCysArgIyLeuArgTyr 60

Db 131 AACTCTTTGGCCAGGCTCTGACTTAAAGCTTACCGGAAGAGATGGCCGGCTCGGATGG 190

QY 61 LeuAsnTyrIleuArgProAspValArgArgIyAAsnIleThrProGluGluGluLeuLeu 80

Db 191 CTAACTACCTCCGCTCTGATGTAGAAAGGGAATATTACACCGGAAGAACGCTTTTG 250

QY 81 IleMetGluLeuHsAlaIySArgIySArgIySArgIySArgIySArgIySArgIySArgIyS 100

Db 251 ATCATGAGACTTCAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 310

QY 101 GYAAGThAspAsnGluIleIySAsnTyrTyrArgThArgIleGluHsIleIyS 120

Db 311 GGAAGAGCTGATTAAGATTAAGAACTAGTGAAGACAGGATCCAGAGAGAGAGAGAGAGAG 370

QY 121 GlnAla---GluAsnPhgGlnGlnGlnSerAAsnSerGluIleAsnAspHsGln 139

Db 371 CAAGCTTCCAGCAGCTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 430

QY 140 AlaSerThrSerHsValSerThrMetAlaGluProMetGluMetTyrSerProProCys 159

Db 431 GCTTGACATGACCAAGTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 490

QY 160 TyrGlnIyMetLeuGluProPheSerThrGlnPheProThrIleAsnPro 176

Db 491 TATCAAGGATGTATGATTCATTTCAATTCAAGTCCACAA---AATCTT 538

RESULT 13

BMS28383 542 bp mRNA linear EST 19-FEB-2002  
sa157ef09.y1 Gm-cl061 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-cl061-3330 5' similar to TR:Q39028 Q39028 ATMYB2. ;, mRNA  
sequence.

ACCESSION BMS28383  
VERSION BMS28383.1 GI:18734780  
KEYWORDS EST.

SOURCE soybean.  
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 542)  
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna  
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers  
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk  
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann  
R., Waterston,R. and Wilson,R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800)-533-4363 or contact: ccu@resgen.com web site:  
www.resgen.com

TITLE JOURNAL  
COMMENT

Seq primer: -40RP from Glibco  
High quality sequence stop: 423.  
Location/Qualifiers

## FEATURES

source

1. .542  
/organism="glycine max"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-cl061-3330"  
/clone\_lib="Gm-cl061"  
/tissue\_type="mature flowers of field grown plants"  
/lab\_host="DH10B"  
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 191 a 121 c 107 g 123 t

ORIGIN

Alignment Scores:

Pred. No.:	7.88e-81	Length:	542
Score:	774.50	Matches:	150
Percent Similarity:	85.71%	Conservative:	6
Best Local Similarity:	82.42%	Mismatches:	21
Query Match:	13	Indels:	5
		Gaps:	3

US-10-021-811-36 (1-206) x BM528383 (1-542)

QY 1 MetAspLysLysGlnGlnCysLeuThrSerGlnAspProGluValArgLysGlyProTyr 20  
Db 6 ATGACACAAAAACCATGCACTCATCTCATCTGACAGAGAAAGCAACCATCG 65  
QY 21 ThrMetGlnGluAspLeuLeuMetAsnTyrIleAlaAsnHisGlyGlnGlyValTyr 40  
Db 66 ACCATGGAAAGACTGATCTGATTAACATATTCGAATTCACGGTGAAGGTGTTGG 125  
QY 41 AsnSerLeuAlaValAlaGlyLeuLysArgAsnGlyLysSerCysArgLysLeuTyr 60  
Db 126 AACCTCTTAAAGCAAGCTCTGCTCAACGAAGGAAAGATTTTCGACCTCCGTGG 185  
QY 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGlnGlnGlnLeu 80  
Db 186 CTTAACTACCTTCTGCTGATGTAGAGAGAAACATTACCCGAGAGAACACTTTTG 245  
QY 81 IleMetGlnLeuHisAlaLysTyrDgLYAsnArgTyrSerLysIleAlaLysHisLeuPro 100  
Db 246 ATCATAGAACTTCAAGCAAGTGGGGCAATAGTGTCTCAAAATTGCAAAACATCTTCCA 305  
QY 101 GlyArgTyrAspAsnGlnIleLysAsnTyrTrrArgTrrArgIleGlnLysHisIleLys 120  
Db 306 GGAAGAACTGCAATGAAATTAGAACTTCGAGAACAGATCCAAAGACATTAAG 365  
QY 121 GlnAlaGluAsnPheGlnGlnGlnSerSerAsnSerGlnIleAsnAspHisGlnAla 140  
Db 366 CAACTGAGACTTCACACCAACAT-----CGTAATTCAGGAATTAATGATCATCAAGCA 419  
QY 141 -----SerThrSerHisValSerThrMetAlaGluProMetGluMetCysSerPro 158  
Db 420 AGCACTAGTACTAGCAAGGTGTCCACCATGCAATCCCAATGACTTCTCTCCACCC 479  
QY 159 CysTyrGlnGlyMetLeuGlnProPheSerThrGlnPheProThrIleAsnProAspGln 178  
Db 480 TCATATACCAAGAACTTTTGAACCATTCACACCTCAATTCCTTACATCACT--GATCAA 536  
QY 179 SerSer 180  
Db 537 TCAAGT 542  
RESULT 14

AI486576 681 bp mRNA linear EST 18-MAY-2001  
LOCUS AI486576  
DEFINITION EST244897 tomato ovary, TMU Lycopersicon esculentum cDNA clone  
CLDB6E18, mRNA sequence.  
ACCESSION AI486576  
VERSION AI486576.1 GI:4381947  
KEYWORDS EST.  
SOURCE tomato.  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.  
REFERENCE  
1 (bases 1 to 681)  
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,  
Liang,F., Upton,J., Roming,C.M., Craven,M.B., Fujii,C.Y., Bowman,  
C.L., Niernm,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tankeley,  
S.D. and Giovannoni,J.  
Generation of ESTs from tomato carpel tissue  
Unpublished (1999)  
COMMENT  
JOURNAL  
Contact: CUGI  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>.  
location/Qualifiers  
1. .681  
/organism="Lycopersicon esculentum"  
/cultivar="T1A96"  
/db\_xref="taxon:4081"  
/clone="CLDB6E18"  
/tissue\_type="carpel"  
/dev\_stage="5 days pre-anthesis to 5 days post-anthesis"  
/lab\_host="X11-Blue MRF"  
/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; cLED - Tomato Carpel EST Library. OligodT-primed and directionally cloned cDNA in vector lambda ZAP II with 5' and 3' ends located at the EcoRI and XhoI sites, respectively."

BASE COUNT 248 a 101 c 145 g 187 t

ORIGIN

Alignment Scores:

Pred. No.:	9.82e-75	Length:	681
Score:	724.00	Matches:	139
Percent Similarity:	80.31%	Conservative:	16
Best Local Similarity:	72.02%	Mismatches:	26
Query Match:	9	Indels:	12
		Gaps:	4

US-10-021-811-36 (1-206) x AI486576 (1-681)

QY 16 ArgLysGlyProTrrThrMetGlnGluAspLeuIleLeuMetAsnTyrIleAlaAsnHis 35  
Db 9 AGGAAAGGCGCTTGAGCTATGGAAGATTTAATTCATTACCTAATGCTATCAT 68  
QY 36 GlnGlnGlnValTrrAsnSerLeuAlaLysAlaGlyLeuLysArgAsnGlyLysSer 55  
Db 69 GGTGAAGGTGTGGAACTCTAGCTAATCTGCTGCTCAACGTAAGTGAAGAACT 128  
QY 56 CysArgLeuArgTrrLeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrPro 75  
Db 129 TGTAGACTCGATGCTTAATTAATTCAGCTGATGTCAAGAGGGGTAAATTCACCT 188  
QY 76 GlnGlnGlnLeuLeuIleMetGlnLeuHisAlaLysTrrDgLYAsnArgTrrSerLysIle 95  
Db 189 GAAAGCAACTTTTGAATTTGGAACCTACATGCTTAAGTGGGAAACAGATGTCAAAAAAT 248  
QY 96 AlaLysHisLeuProGlnValArgTrrAspAsnGlnIleLysAsnTyrTrrArgTrrArgIle 115  
Db 249 GCGAAGCAATTTGCTTGAAGAACGATACGATTAACAACTACTGAGAGCTAGATT 308

